

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 28, 2003, 03:48:38 ; Search time 10747 Seconds
(without alignments)
11760.816 Million cell updates/sec

Title: US-09-730-559B-7

Perfect score: 4343

Sequence: 1.atggagccccaagaagcagg.....attcacagcagtcacctctt 4343

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_em.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_rnan.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
C 1	3301.2	76.0	169377	9	AC027228	Homo sapi
C 2	1420.4	32.7	1546	9	BC017880	Homo sapi
C 3	1023	23.6	2389	9	HUMIREBPA	Human iron-
C 4	1023	23.6	3280	6	AX400060	Sequence
C 5	990	22.8	2867	6	AX377520	Sequence
C 6	990	22.8	2867	9	HSU20180	Human iron-
C 7	944.2	21.7	3341	9	AK027033	Homo sapi
C 8	877.2	20.2	3770	6	AX377522	Sequence
C 9	877.2	20.2	3770	10	RNU20181	U20181 Rattus norv
C 10	396.2	9.1	309662	2	AC004469	Homo sapi
C 11	385.4	8.9	176314	2	AC044886	Homo sapi
C 12	381.2	8.8	141037	9	AC096711	Homo sapi
C 13	380.2	8.8	183658	2	AC032021	Homo sapi
C 14	380.2	8.8	186131	9	AL158192	Human DNA
C 15	377.6	8.7	170881	2	AC122718	Homo sapi
C 16	377.2	8.7	92702	9	AC091829	Homo sapi
C 17	374.6	8.6	206454	9	AC018737	Homo sapi
C 18	373.6	8.6	110000	2	AL390202_05	Continuation (6 of
C 19	373.6	8.6	162609	9	AC108022	Homo sapi
C 20	372.2	8.6	40182	9	AC092824	Homo sapi
C 21	372.2	8.6	159385	9	AC007619	Homo sapi
C 22	371	8.5	155862	9	AC018464	Homo sapi
C 23	370.8	8.5	186321	9	AC006511	Homo sapi
C 24	369	8.5	170330	9	AC092896	Homo sapi
C 25	367.6	8.5	102837	9	HS52202	Z95329 Human DNA s
C 26	367.2	8.5	177647	9	AC091969	Homo sapi
C 27	366.8	8.4	99781	9	AF031078	Homo sapi
C 28	366.8	8.4	112756	9	AF030876	Homo sapi
C 29	364.4	8.4	195342	9	AC009120	Homo sapi
C 30	364	8.4	166032	2	AC023645	Homo sapi
C 31	363.6	8.4	86441	2	AL390762	Homo sapi
C 32	363.6	8.4	100000	9	AP000519	Homo sapi
C 33	363.6	8.4	162227	9	AB023056	Homo sapi
C 34	363.4	8.4	156399	9	AC002456	Homo sapi
C 35	361.8	8.3	175789	2	AP000425	Homo sapi
C 36	361.6	8.3	21896	9	AL591364	Human DNA
C 37	360.6	8.3	153221	2	AC067998	Homo sapi
C 38	360	8.3	134026	2	AC069317	Homo sapi
C 39	359.4	8.3	131928	9	AC091738	Homo sapi
C 40	358.4	8.3	219357	9	AC107464	Homo sapi
C 41	358.2	8.2	160302	2	AC129730	Homo sapi
C 42	358.2	8.2	193970	2	AL357119	Homo sapi
C 43	356.6	8.2	116400	9	AC055752	Homo sapi
C 44	356.6	8.2	152528	2	AC016063	Homo sapi
C 45	355.6	8.2	162232	9	AC094095	Homo sapi

ALIGNMENTS

RESULT 1	AC027228/c	AC027228	169377 bp	DNA	linear	PRI 06-AUG-2002
LOCUS	AC027228/c	Homo sapiens chromosome 15, clone RP11-650L12, complete sequence.				
DEFINITION	AC027228	HTG.				
ACCESSION	AC027228.16	GI:22024598				
VERSION	HTG.					
KEYWORDS	human.					
SOURCE	Homo sapiens					
ORGANISM	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1 (bases 1 to 169377)					
AUTHORS	Birren,B., Nusbaum,C. and Lander,E.					
TITLE	Homo sapiens chromosome 15, clone RP11-650L12					
JOURNAL	Unpublished					

REFERENCE
AUTHORS

2 (bases 1 to 169377)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,A., Allen,N., Anderson,S., Baldwin,J., Barna,N., Bastien,V., Bloom,T., Boukhvaliy,L., Boukhvalter,B., Boukhvaliy,L., Chang,J., Chazaro,B., Choepel,Y., Collamore,A., Campiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Dodge,S., Doming,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Pierre,N., Pisanic,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and Zody,M.

Direct Submission
 Submitted (28-MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

TITLE
JOURNALREFERENCE
AUTHORS

3 (bases 1 to 169377)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boukhvaliy,L., Boukhvalter,B., Brown,A., Camarata,J., Campiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzGerald,M., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Straus,K., Travers,M., Travis,N., Trigilio,J., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
 Submitted (12-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

TITLE
JOURNALREFERENCE
AUTHORS

4 (bases 1 to 169377)
 Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boukhvaliy,L., Boukhvalter,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collamore,A., Cooke,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupack,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
 Submitted (31-JUL-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE
AUTHORS

5 (bases 1 to 169377)
 Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boukhvaliy,L., Boukhvalter,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collamore,A., Cooke,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupack,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
 Submitted (06-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jul 31, 2002 this sequence version replaced gi:21392498.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

TITLE
JOURNAL

COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L8639

Center clone name: 650_L_12

FEATURES
source

----- Location/Qualifiers
 1. .169377
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="15"
 /map="15"
 /clone="RP11-650L12"
 /clone.lib="RPC1-11 Human Male BAC"
 complement(983..1324)
 /rpt_family="MLT1A1"
 complement(1338..1562)
 /rpt_family="MER20"
 1750..1771
 /rpt_family="AT-rich"
 2162..2206
 /rpt_family="L3"
 complement(2334..2619)
 /rpt_family="Alusq"
 complement(2972..3255)
 /rpt_family="Alusp"
 complement(4684..4972)
 /rpt_family="Alusg"
 5215..5301
 /rpt_family="L1PB1"
 complement(5623..5926)
 /rpt_family="Alusx"
 8353..8444
 /rpt_family="MIR"
 9364..9808
 /rpt_family="L2"
 9749..9760
 /note="<30 qual SNGL region"
 9784..9794 qual SNGL region"
 /note="<30 qual SNGL region"
 9809..10055
 /rpt_family="THE1B"
 10058..10147
 /rpt_family="L2"

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region	10371..10394	/rpt_family="Alu_rich"	
repeat_region	complement(10632..10817)		
repeat_region	/rpt_family="MER104"		
repeat_region	10921..11241		
repeat_region	/rpt_family="AluSx"		
repeat_region	11800..12061		
repeat_region	/rpt_family="AluSg"		
repeat_region	complement(12062..12172)		
repeat_region	/rpt_family="MIR3"		
repeat_region	complement(13299..13602)		
repeat_region	/rpt_family="AluY"		
repeat_region	complement(13832..13896)		
repeat_region	/rpt_family="L2"		
repeat_region	14264..14369		
repeat_region	/rpt_family="L2"		
repeat_region	complement(14504..14810)		
repeat_region	/rpt_family="AluSx"		
repeat_region	14904..15380		
repeat_region	/rpt_family="L1ME4A"		
repeat_region	15921..16013		
repeat_region	/rpt_family="L2"		
repeat_region	16275..16381		
repeat_region	/rpt_family="AluSg/x"		
repeat_region	complement(16712..17016)		
repeat_region	/rpt_family="AluSx"		
repeat_region	complement(17131..17377)		
repeat_region	/rpt_family="L2"		
repeat_region	17469..17791		
repeat_region	/rpt_family="AluSx"		
repeat_region	complement(17816..17938)		
repeat_region	/rpt_family="MER45R"		
repeat_region	complement(18485..18787)		
repeat_region	/rpt_family="AluSx"		
repeat_region	complement(19073..19363)		
repeat_region	/rpt_family="AluYb"		
repeat_region	complement(19939..20038)		
repeat_region	/rpt_family="L2"		
repeat_region	complement(20193..20504)		
repeat_region	/rpt_family="AluJo"		
repeat_region	20569..20876		
Query Match	76.0%;	Score 3301.2;	DB 9; Length 169377;
Best Local Similarity	98.6%;	Pred. No. 0;	
Matches 3431;	Conservative	3; Mismatches	31; Indels 13; Gaps 10;
QY	872	TTCTGCGGTGGGGGTTGGAGGCAATGGAACAGACAGAGTTATGCTTGGTCTGCCAGTTT	931
Db	123784	TTATTAGTTAGGGTTGGAGGCAATGGAACAGACAGAGTTATGCTTGGTCTGCCAGTTT	123725
QY	932	CTCTTACTTTACCAGAGGTGGTGGATGGATGTAAGTAACTGGGTCTATCAAAACCCCTTTTGTTA	991
Db	123724	CTCTTACTTTACCAGAGGTGGTGGATGGATGTAAGTAACTGGGTCTATCAAAACCCCTTTTGTTA	123665
QY	992	CATCCATAGATGTTGTTCTGGTATTTACAAAGTAAGTTAAAGTTGGTGGTCTCATGA	1051
Db	123664	CATCCATAGATGTTGTTCTGGTATTTACAAAGTAAGTTAAAGTTGGTGGTCTCATGA	123605
QY	1052	CTTACTGAACATATTTTTATAAAATTTGAAGACTTATGAGACAGAGGATTTGGGTTTC	1111
Db	123604	CTTACTGAACATATTTTTATAAAATTTGAAGACTTATGAGACAGAGGATTTGGGTTTC	123545
QY	1112	ATTACTGCATCTCAGGTCTCTTGGAGCTTAGCCACATCATCATGATATATAGTAATAA	1171
Db	123544	ATTACTGCATCTCAGGTCTCTTGGAGCTTAGCCACATCATCATGATATATAGTAATAA	123485
QY	1172	CAACAAACAGACATTTAGTTTGTCTACTATAATAACAAAGAAATTTGGTGGTCTCACTTA	1231
Db	123484	CAACAAACAGACATTTAGTTTGTCTACTATAATAACAAAGAAATTTGGTGGTCTCACTTA	123425
QY	1232	TGTTAGTCTATTTAGTCCCTTATAACAGCCTGTGAGATGGATATATCTATTTCTCAT	1291
Db	123424	TGTTAGTCTATTTAGTCCCTTATAACAGCCTGTGAGATGGATATATCTATTTCTCAT	123365

QY	1292	GTAACCTCTGAGAAACTAAAGGTACAGTAGGGTTTAGTGACTTACCAAAAGGTCGAAGGCC	1351
Db	123364	GTAACCTCTGAGAAACTAAAGGTACAGTAGGGTTTAGTGACTTACCAAAAGGTCGAAGGCC	123305
QY	1352	TGAGTA-TAAGGGTAGAGCAAGATTCCAGGCAGTCAGATTCCTGAGTCATGTT---CTA	1407
Db	123304	TGAGTAGTAAGGGTAGAGCAAGATTCCAGGCAGTCAGATTCCTGAGTCATGTTGCTTTA	123245
QY	1408	ACCATTATGCCTTATTAGTGCCTTTGCTTAAATAACACATTCGCTGGACTACATATTTT	1467
Db	123244	ACCATTATGCCTTATTAGTGCCTTTGCTTAAATAACACATTCGCTGGACTACATATTTT	123185
QY	1468	TTTTCTCTTTTAACTTGAATAAAAAATGTTTAGCAAAAGTTGAWTGTGCTCT	1527
Db	123184	TTTTCTCTTTTAACTTGAATAAAAAATGTTTAGCAAAAGTTGAWTGTGCTCT	123126
QY	1528	TTAATTAATTAATTTGCCCGTTAGAAACGTTGCTCTACTAAGTAATGCTTTCAAAAACA	1587
Db	123125	TTAATTAATTAATTTGCCCGTTAGAAACGTTGCTCTACTAAGTAATGCTTTCAAAAACA	123066
QY	1588	TGGACTGTAGAAATGTATATATCATTTTCTGTTGCCCTTTTAAACATTTCTCTCGATTA	1647
Db	123065	TGGACTGTAGAAATGTATATATCATTTTCTGTTGCCCTTTTAAACATTTCTCTCGATTA	123006
QY	1648	TTATGTAATAAACTCTCTCTCTCAATTTTAAATACTGGCTTCAGAACTTCAATACATAC	1707
Db	123005	TTATGTAATAAACTCTCTCTCTCAATTTTAAATACTGGCTTCAGAACTTCAATACATAC	122946
QY	1708	ACTGAGCTTTGTTAAGCATATTAATACACAGCTCAGGATTTTCCCTAGTGAACAATAATT	1767
Db	122945	ACTGAGCTTTGTTAAGCATATTAATACACAGCTCAGGATTTTCCCTAGTGAACAATAATT	122886
QY	1768	GTAACCTCTTCTCTCTCTCTGCTTGTCTGCTTGTCTTAACTTAACTTAACTTAACTT	1827
Db	122885	GTAACCTCTTCTCTCTCTCTGCTTGTCTGCTTGTCTTAACTTAACTTAACTTAACTT	122826
QY	1828	TTTGTGTAATGAATGTACCTGGAATAATGTTCCACATATAATTTCCAAATTTGAGTCCCAATC	1887
Db	122825	TTTGTGTAATGAATGTACCTGGAATAATGTTCCACATATAATTTCCAAATTTGAGTCCCAATC	122766
QY	1888	TCAGCATTTTGGTTAGATTTATGTTAGCAAGCTTTCTGGATATCTCCAGTCTAAGGAAA	1947
Db	122765	TCAGCATTTTGGTTAGATTTATGTTAGCAAGCTTTCTGGATATCTCCAGTCTAAGGAAA	122706
QY	1948	TGATTAATGCTCTCTCTCAGCATTTGGTATGATCCTTCTTCCCTTAACTAGAAAAAAT	2007
Db	122705	TGATTAATGCTCTCTCTCAGCATTTGGTATGATCCTTCTTCCCTTAACTAGAAAAAAT	122646
QY	2008	TGGCATCTTTAGAGAAATATTGATTCAAGGTATGATACCAAAAGATCAAGTAGTAATAATG	2067
Db	122645	TGGCATCTTTAGAGAAATATTGATTCAAGGTATGATACCAAAAGATCAAGTAGTAATAATG	122586
QY	2068	GGAAATGACAGATTAATCCTAGAGGAAAGGAGTATCCCATTTATGTTTACAGAAATCA	2127
Db	122585	GGAAATGACAGATTAATCCTAGAGGAAAGGAGTATCCCATTTATGTTTACAGAAATCA	122526
QY	2128	ATTCTTTACTTTAGACATCTCGAAACCTAACGCTGCTTTTAGCCTTCTCTAGCTGTTTT	2187
Db	122525	ATTCTTTACTTTAGACATCTCGAAACCTAACGCTGCTTTTAGCCTTCTCTAGCTGTTTT	122466
QY	2188	TTCCCTGACAATATTACTGTGTTTGTGTTTACATTTTACGTTTAACTTAACTTAACTT	2247
Db	122465	TTCCCTGACAATATTACTGTGTTTGTGTTTACATTTTACGTTTAACTTAACTTAACTT	122406
QY	2248	ATTATATATGTTTACATTTTATGAATATATGATTACTTCTTTTGTAGATCCCTGTTCCA	2307
Db	122405	ATTATATATGTTTACATTTTATGAATATATGATTACTTCTTTTGTAGATCCCTGTTCCA	122347
QY	2308	TTTGTGATCCCTTATAGGAATTAATCCTGTAATGTTTTTTTGTATGAGACAGCATTTGGTT	2367
Db	122346	TTTGTGATCCCTTATAGGAATTAATCCTGTAATGTTTTTTTGTATGAGACAGCATTTGGTT	122287

QY 2368 GTAATATCTAATCTGTGTTCTTCTATCCTAAAAATAAACCATAGCCGGCGCGGTG 2427
|||||
Db 122286 GTAATATCTAATCTGTGTTCTTCTATCCTAAAAATAAACCATAGCCGGCGCGGTG 122227
|||||
QY 2428 GCTACGCCGTGTAATCCAGACACTTTGGAGCGCGAGCGGGTGGATCATG#SGTCAAGGA 2487
|||||
Db 122226 GCTACGCCGTGTAATCCAGACACTTTGGAGCGCGAGCGGGTGGATCATGAGTCAAGGA 122167
|||||
QY 2488 GATCAGACCATCTCGGTGTAACAAGGTCAAAACCCGCTCTCTACTATAAATAACAAAAATT 2547
|||||
Db 122166 GATCAGACCATCTCGGTGTAACAAGGTGAAAACCCCGCTCTCTACTATAAATAACAAAAATT 122107
|||||
QY 2548 AGCCGGCGCGGTGCGCGCGCGCTGTAGTCCAGCTTCTCGGGAGGCTGAGGCGAGAGAA 2607
|||||
Db 122106 AGCCGGCGCGGTGCGCGCGCGCTGTAGTCCAGCTTCTCGGGAGGCTGAGGCGAGAGAA 122047
|||||
QY 2608 TGGCGTGAAACCGGGAGCGGAGCTTGCAGTGAAGCCAAAGATTGGCCCACTGCATCGCGCA 2667
|||||
Db 122046 TGGCGTGAAACCGGGAGCGGAGCTTGCAGTGAAGCCAAAGATTGGCCCACTGCATCGCGCA 121987
|||||
QY 2668 GTCCGGCTGGCGGACAGCGGAGCTCTGCTCAAAAAAATAAAAAAATAAAAAAAC 2727
|||||
Db 121986 GTCCGGCTGGCGGACAGCGGAGCTCTGCTCAAAAAAATAAAAAAATAAAAAAAC 121927
|||||
QY 2728 CATAAATGAGGAACCGCATCTTTACACTTAGGGTTTGAGTTTCTGTATCTATAAAAAAGG 2787
|||||
Db 121926 CATAAATGAGGAAC--CATCTTTACACTTAGGGTTTGAGTTTCTGTATCTATAAAAAAGG 121869
|||||
QY 2788 GTTGTGATTAAGTATCCCTGGCACTTATAAATGTTAGGCTTAATAATTATTATAGAT 2847
|||||
Db 121868 GTTGTGATTAAGTATCCCTGGCACTTATAAATGTTAGGCTTAATAATTATTATAGAT 121809
|||||
QY 2848 CGAGGATAGTTTCATCTTAGTCGCTCCTTAGTCACCTTCCTATACCAATCTGAGACC 2907
|||||
Db 121808 CGAGGATAGTTTCATCTTAGTCGCTCCTTAGTCACCTTCCTATACCAATCTGAGACC 121749
|||||
QY 2908 ATTTTACAAATTTAGAAAACAGCAAAATACTGGTTGGGTTTACTGTAGTATATAAACCAAG 2967
|||||
Db 121748 ATTTTACAAATTTAGAAAACAGCAAAATACTGGTTGGGTTTACTGTAGTATATAAACCAAG 121689
|||||
QY 2968 AAAATAAATTTTAGAGGAATTAAGTTTGAACACCATGTTAAACAAATCTTACCAAAAGTG 3027
|||||
Db 121688 AAAATAAATTTTAGAGGAATTAAGTTTGAACACCATGTTAAACAAATCTTACCAAAAGTG 121629
|||||
QY 3028 GGATTTGCCGTGATTAAGATGCTGTAACCATTTGGGCCAGTAGTTATAATTTGAAAAA 3087
|||||
Db 121628 GGATTTGCCGTGATTAAGATGCTGTAACCATTTGGGCCAGTAGTTATAATTTGAAAAA 121569
|||||
QY 3088 TGTATATAGCCAATATATAATTTTATTTAATATACAGTTTCATCAGTCTATAGTAT 3147
|||||
Db 121568 TGTATATAGCCAATATATAATTTTATTTAATATACAGTTTCATCAGTCTATAGTAT 121509
|||||
QY 3148 TTCAATTAAGTCTAAGATGCCATGCTGGTTAGCAACACCACTGTTTATGCACTGCTAA 3207
|||||
Db 121508 TTCAATTAAGTCTAAGATGCCATGCTGGTTAGCAACACCACTGTTTATGCACTGCTAA 121449
|||||
QY 3208 GAAAGAATAAAGGGCTGTGTGCAGTGGCTCACACCTGTGGGACGCCAAGGCGAGGACATC 3267
|||||
Db 121448 GAAAGAATAAAGGGCTGTGTGCAGTGGCTCACACCTGTGGGACGCCAAGGCGAGGACATC 121389
|||||
QY 3268 ACTTGAGCCAGCAAGTTCAAGACCAACCTGGTCAACATTTGAAGACCTGCTCTACAAA 3327
|||||
Db 121388 ACTTGAGCCAGCAAGTTCAAGACCAACCTGGTCAACATTTGAAGACCTGCTCTACAAA 121329
|||||
QY 3328 AAAAAAAGTTAAAAAATAGCTGGGTGGCGGTGGCACATGCTGTAGTTCCAGCTACTCT 3387
|||||
Db 121328 AAAAAAAGTTAAAAAATAGCTGGGTGGCGGTGGCACATGCTGTAGTTCCAGCTACTCT 121269
|||||
QY 3388 GGAGGCTAAGGTGGGAGGATTTGTAGAGCCAGCGGTGGGAAGCTGCAATGAGCTGTGAC 3447
|||||
Db 121268 GGAGGCTAAGGTGGGAGGATTTGTAGAGCCAGCGGTGGGAAGCTGCAATGAGCTGTGAC 121209
|||||
QY 3448 CACACCACCTCCCGCTCCAGCGTGGGCACAGATGAGACCCCTGTTCTTAAAAAGAAAAAG 3507
|||||

Db 121208 CACACACCTGCGCTCCAGCCTGGGCAACAGAGTGAGACCTGTTTCTAAAAAGAAAG 121149
|||||
QY 3508 AAAAAAGGCTGCGACCTAAACAGACACACTATTGAGTTGAGGTACCCCTGATTTCAAAGA 3567
|||||
Db 121148 -AAAAAGGCTGCGACCTAAACAGACACACTATTGAGTTGAGGTACCCCTGATTTCAAAGA 121090
|||||
QY 3568 CATGAAAAATGTTAATATATAGCCACCTTGAAGCTTTTCAGSGCCCTTTCTACCCCTGAATTA 3627
|||||
Db 121089 CATGAAAAATGTTAATATATAGCCACCTTG-AGCTTTTCAGGCCCTTTCTACCCCTG-ATTAA 121032
|||||
QY 3628 CAGTGACATTGGACAGTCTTCTCTTTACTTCTTATCTTTAAATACCCCAAAACCAGAA 3687
|||||
Db 121031 CAGTGACATTGGACAGTCTTCTCTTTACTTCTTATCTTTAAATACCCCAAAACCAGAA 120972
|||||
QY 3688 TGAGTTGATTCATAAGGACAAATGAAGGATCTCATCTCTTACCATCACTAGTATTGGTTA 3747
|||||
Db 120971 TGAGTTGATTCATAAGGACAAATGAAGGATCTCATCTCTTACCATCACTAGTATTGGTTA 120912
|||||
QY 3748 AAAATTTTATTTTATAGTTTTCAGACAAATCGTTGCTAAATCTTATCTTTCGAAATTTGTAT 3807
|||||
Db 120911 AAAATTTTATTTTATAGTTTTCAGACAAATCGTTGCTAAATCTTATCTTTCGAAATTTGTAT 120852
|||||
QY 3808 GTGTTCTGTGATTCCTTATATAGCACCTCAGGCAAGTAGGAGTGGCTGGAAAGTTTGT 3867
|||||
Db 120851 GTGTTCTGTGATTCCTTATATAGCACCTCAGGCAAGTAGGAGTGGCTGGAAAGTTTGT 120792
|||||
QY 3868 TGAGTTTGTGGAAGTGGAGTTTCACAATATCTATAGTTGATCGAACCTACATAGCAAA 3927
|||||
Db 120791 TGAGTTTGTGGAAGTGGAGTTTCACAATATCTATAGTTGATCGAACCTACATAGCAAA 120732
|||||
QY 3928 CATGTGTCGGAATATGTTGCTATCTCAGCTTTTCCCTGTTGACAATGTGACATTAA 3987
|||||
Db 120731 CATGTGTCGGAATATGTTGCTATCTCAGCTTTTCCCTGTTGACAATGTGACATTAA 120672
|||||
QY 3988 ACATTTAGACATACAGTGAAGATAGAAGATAGAAGATCACTAGATAAACAATGTTACATTTCC 4047
|||||
Db 120671 ACATTTAGACATACAGTGAAGATAGAAGATAGAAGATCACTAGATAAACAATGTTACATTTCC 120612
|||||
QY 4048 AATGTGTTGATATATTTTATAAATTTACTACCTTATCCATGTTATTTACTACTACCAAA 4107
|||||
Db 120611 AATGTGTTGATATATTTTATAAATTTACTACCTTATCCATGTTATTTACTATTTCAAAA 120552
|||||
QY 4108 ATTACATTTGTTGAACAACAACCTTTCAAGCAACATCAGATGCTTTTAAAGAGTGTG 4167
|||||
Db 120551 ATTACATTTGTTGAACAACAACCTTTCAAGCAACATCAGATGCTTTTAAAGAGTGTG 120492
|||||
QY 4168 TGTCTCAAAACCCCTAGTTCCTGTGACACATTTGAAAGCAATTTAAAGGAATTTTCAAAAC 4227
|||||
Db 120491 TGTCTCAAAACCCCTAGTTCCTGTGACACATTTGAAAGCAATTTAAAGGAATTTTCAAAAC 120432
|||||
QY 4228 CATTGATCCTGACTGACTGTTT-CCCATTAATGATGATACCTCCCTCTACTTAGGGG 4286
|||||
Db 120431 CATTGATCCTGACTGACTGTTTCCCATTAATGATGATACCTCCCTCTACTTAGGGG 120372
|||||
QY 4287 TCATAGG-TGCAATTTAATGGAGTCAAGCTTAAACATATTTACAGCAGTCCCTCTCT 4343
|||||
Db 120371 TCATAGGTTGCAATTTAATGGAGTCAAGCTTAAACATATTTACAGCAGTCCCTCTCT 120314
|||||

RESULT 2

BC017880 1546 bp mRNA linear PRI 06-DEC-2001
LOCUS Homo sapiens, clone MGC:22754 IMAGE:4277855, mRNA, complete cds.
DEFINITION BC017880
ACCESSION BC017880
VERSION BC017880.1 GI:17389723
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (bases 1 to 1546)
AUTHORS Strausberg, R.

Db 1435 AGGCAGTCAGATTCCTGAGTCCATTGCTTAAACCAATTATGCTTATTAGTGGCTTGTC 1494

QY 1437 CTTAATAAACACTTCTGCTG 1454

Db 1495 CTTAATAAACACTTCTGCTG 1512

RESULT 3

HUMIREBPA

LOCUS

DEFINITION Human iron-responsive element-binding protein/iron regulatory protein 2 (IRE-BP2/IRP2) mRNA, partial cds.

ACCESSION M58511.1

VERSION M58511.1 GI:897825

KEYWORDS RNA regulatory protein; iron-responsive element-binding protein.

SOURCE Human cDNA to mRNA.

ORGANISM Homo sapiens

REFERENCE

AUTHORS Rouault, T.A., Tang, C.K., Kaptain, S., Burgess, W.H., Haile, D.J., Samanigo, F., McBride, O.W., Harford, J.B. and Klausner, R.D.

TITLE Cloning of the cDNA encoding an RNA regulatory protein--the human iron-responsive element-binding protein

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 87 (20), 7958-7962 (1990)

MEDLINE 91045916

PUBMED 2172968

REFERENCE

AUTHORS Samanigo, F., Chin, J., Iwai, K., Rouault, T.A. and Klausner, R.D.

TITLE Molecular characterization of a second iron-responsive element-binding protein, iron regulatory protein 2. Structure, function, and post-translational regulation

JOURNAL J. Biol. Chem. 269 (49), 30904-30910 (1994)

MEDLINE 95074121

PUBMED 7983023

REFERENCE

AUTHORS Rouault, T.A., Haile, D.J., Downey, W.E., Philpott, C.C., Tang, C., Samanigo, F., Chin, J., Paul, R., Orloff, D., Harford, J.B. et al.

TITLE An iron-sulfur cluster plays a novel regulatory role in the iron-responsive element binding protein

JOURNAL Biometals 5 (3), 131-140 (1992)

MEDLINE 93044098

PUBMED 1421965

REFERENCE

AUTHORS Rouault, T.A.

TITLE Direct Submission

JOURNAL Submitted (17-DEC-1990) Tracey A. Rouault, Cell Biology and Metabolism Branch, NICHD, NIH, Bethesda, MD 20992, USA

COMMENT On Jul 14, 1995 this sequence version replaced gi:186547.

FEATURES

source

1..2889

/organism="Homo sapiens"

/db_xref="taxon:9606"

/map="9p"

/cell_line="YT"

/cell_type="T cell"

/dev_stage="adult"

1..2889

/gene="IRE-BP2/IRP2"

1..>2889

/gene="IRE-BP2/IRP2"

/codon_start=1

/product="iron-responsive element-binding protein/iron regulatory protein 2"

/protein_id="AA69901.1"

/db_xref="GI:897825"

/translation="MDAPKAGAFEYLIELNDSSHKFFDVSGLGTYDVLPSYIRV LLEAARNCQGLMKEDVMNILDWTKOSNVPEFFPARVLQDFTGIPAMVDFAAM REAVTLGGPEKHVPACPTDLVDHSLQIDFSKCAIONAPNGGDLQKAGLSPLK VQPKLFCGGTTCRSCDSGELGRNSGTFSSQIENTPILCFHPQVPPEFVLKNO VEPGRNRERLQFKWSSRVLKNVPIPPGTGMAHQINLEYLSRVVVFEEKDLFPDSV VGTDSHTMTVMNGILGWGGIETEAVMGLPVLTPPEVVGCELTGSSNPFVTSID VVLGITTHLRQVGVAGKVFVEFFGSGVSQLSIVDRTTIANMCPYGAILLSFFPVDNVL

KHLEHTGFSKAKLESMEYLYKAVKLFNRDNQSSGEPEYSQVIOINLSIVPSVSGPKR
 PODRVAVTDMKSDFOACLNKGVKGFQIAAEKQKDIVSIHYGESEYKLSHGSVVIAA
 VLSCTNNPNPSVMLAAGLLAKKAVEAGLVKVPVIRTSLSGSGMWHVLSLSSGVLPYL
 SKLGFELVGYGCTCVCNCTAPLSDAVINAVKOGDLVTCGILSGNKFGRUCDCVRAN
 YLASPPLVAYATAGTVNIDFOTPLGTDPKNIYLDHIMPREEVHRVEEHVILS
 MKALKDKIEMGNKRNLSLEAPDSVLFPMWLKSTYIRCPSPFDKLTKEPZALQAIENA
 HVLVLDGDSVTTHDISPAGSIARNSAAKLTNRGLTPREFNSYKLRGNDADNATRTGT
 FANIKLFNKGFKVAPAKTHIFPSGQTLDFEAEALYQKEGIPLIILAKKKYGSNSRD
 WAAKGPYLLGVKAVLAEYKIKHDLHIGLIGIAPLOFLFGENADSLSGHRETFSLTF
 PEELSPGTTLNIOQTSGKVFSTASPEDDVEITLYKHGGLLNFVARKFS"

BASE COUNT 879 a 534 c 637 g 839 t

ORIGIN

Query Match 23.6%; Score 1023; DB 9; Length 2889;
 Best Local Similarity 100.0%; Pred. No. 6.8e-205;
 Matches 1023; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGACGCCCCAAAAGCAGGATACGCCCTTGGATGACCTTATTGAAACATTAAATGACAGT 60
 DB 1 ATGGACGCCCCAAAAGCAGGATACGCCCTTGGATGACCTTATTGAAACATTAAATGACAGT 60

QY 61 TCACATAAGAAGTTCTTCGATGATCTAAACTTGGACCAAGTATGATGTTGCGCTTAC 120
 DB 61 TCACATAAGAAGTTCTTCGATGATCTAAACTTGGACCAAGTATGATGTTGCGCTTAC 120

QY 121 TCATAACGGGTCTTGTGGAGGCTGCTGACGAAATTTGATGGCTTTTAAATGAAGAAG 180
 DB 121 TCATAACGGGTCTTGTGGAGGCTGCTGACGAAATTTGATGGCTTTTAAATGAAGAAG 180

QY 181 GAAGATGTTATGAACATTTTAGACTGGAAACCAACAAAGCAATGTTGAAGTGCCTTT 240
 DB 181 GAAGATGTTATGAACATTTTAGACTGGAAACCAACAAAGCAATGTTGAAGTGCCTTT 240

QY 241 TTCCCTGCCCGTGTCTTCTTCAAGATTTTACTGSAATACACCAATGGTGGATTTTGGT 300
 DB 241 TTCCCTGCCCGTGTCTTCTTCAAGATTTTACTGSAATACACCAATGGTGGATTTTGGT 300

QY 301 GCTATGAGGAGGCGAGTGAACACTCTTGGAGTGATCTCGAGAGCCATCTCCACCTTAA 360
 DB 301 GCTATGAGGAGGCGAGTGAACACTCTTGGAGTGATCTCGAGAGCCATCTCCACCTTAA 360

QY 361 CCGACAGATCTTACAGTTGACCATTTTACAAATTTGACTTCAGTAAATGTCAATACAG 420
 DB 361 CCGACAGATCTTACAGTTGACCATTTTACAAATTTGACTTCAGTAAATGTCAATACAG 420

QY 421 AATGCACCAAACTCGAGGTGCTGCACCTGCAGAACGAGAAAGCTCTCCACCTTAA 480
 DB 421 AATGCACCAAACTCGAGGTGCTGCACCTGCAGAACGAGAAAGCTCTCCACCTTAA 480

QY 481 GTGACGCTTAAGAAGCTTCCCTGACAGAGCCAGACTACCTGCCGAGGATCTTGTGATTC 540
 DB 481 GTGACGCTTAAGAAGCTTCCCTGACAGAGCCAGACTACCTGCCGAGGATCTTGTGATTC 540

QY 541 GGAGAACTAGGCGGAACCTCAGAACATTTTTCGCGAGATTGAGAAATACACCACTCCTG 600
 DB 541 GGAGAACTAGGCGGAACCTCAGAACATTTTTCGCGAGATTGAGAAATACACCACTCCTG 600

QY 601 TGTCTCTTTTCATTTCACACCAAGTGCCTGAACTGAAACAGTGTGTTAAAAAATCAAGAAGTA 660
 DB 601 TGTCTCTTTTCATTTCACACCAAGTGCCTGAACTGAAACAGTGTGTTAAAAAATCAAGAAGTA 660

QY 661 GAATTCGGCAGAAATCGAGAGAGGCTTCAGTTTTTAAAGTGGAGTTCAAGAGTTTAAAG 720
 DB 661 GAATTCGGCAGAAATCGAGAGAGGCTTCAGTTTTTAAAGTGGAGTTCAAGAGTTTAAAG 720

QY 721 AATGTGGCAGTATCCCTCCTGGAAGTGAATGGCTCATCAATAAATTAAGTATTTT 780
 DB 721 AATGTGGCAGTATCCCTCCTGGAAGTGAATGGCTCATCAATAAATTAAGTATTTT 780

QY 781 TCAAGAGTGGTTTTTGAAGAAAAAGACCTCTCTCCAGACAGTGTAGTCCGCACAGAT 840
 DB 781 TCAAGAGTGGTTTTTGAAGAAAAAGACCTCTCTCTCCAGACAGTGTAGTCCGCACAGAT 840

QY 841 TCACACATAACAGGATGAATGGTTAGGATCTGGGGTGGGGTGGAGGCATTGAA 900
|||||
Db 841 TCACACATAACAGGATGAATGGTTAGGATCTGGGGTGGGGTGGAGGCATTGAA 900
|||||
QY 901 ACAGAACAGTATAGCTTGGTCTGCCAGTTTCTCTTACTTTACCAGAGGTGGTTGGATG 960
|||||
Db 901 ACAGAACAGTATAGCTTGGTCTGCCAGTTTCTCTTACTTTACCAGAGGTGGTTGGATG 960
|||||
QY 961 GAGTTAACTGGGTGATCAAAACCCCTTTGTTACATCCATAGATGTTGTTCTTGGTATTACA 1020
|||||
Db 961 GAGTTAACTGGGTGATCAAAACCCCTTTGTTACATCCATAGATGTTGTTCTTGGTATTACA 1020
|||||
QY 1021 AAG 1023
|||
Db 1021 AAG 1023
|||

RESULT 4
AX400060 3280 bp DNA linear PAT 06-JUN-2002
LOCUS AX400060
DEFINITION Sequence 231 from Patent WO0218424.
ACCESSION AX400060
VERSION AX400060.1 GI:21336471
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Tang,Y.T., Asundi,V., Zhou,P., Xue,A.J., Ren,F., Zhang,J.,
Wang,J.R., Zhao,Q.A., Wang,D., Liu,C., Drmanac,R.T. and Wehrman,T.
Nucleic acids and polypeptides
Patent: WO 0218424-A 231 07-MAR-2002;
JOURNAL HYSEQ, INC. (US)
FEATURES
source Location/Qualifiers
1..3280
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..2892
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD33504.1"
/db_xref="GI:21336472"
/translation="MDAPKAGYAFEYLIELNDSSHKKFFDVSKLGTKYDVLPSYIRV
LLEAAVRNCDGELMKKEDVMNILDWKTKDSNVEVFPFARVLLQDFTGIPAMVDFPAA
REAVKTLGGDPEKVPACPTDLTVDSLOIDFSCAIONAPNPGGDLQKAGKLSPLK
VOPKLPGRGOTTGRGSCDSGBELRNSGTFFSQIENTPLICFPHLOPVPETVLKNQ
EVEFGNRELROFFKWSRVLNKNAVIPETGMARQINLEYLSRVVFEKDLFFPDSV
VGTDSHTIWNVLGILGWGGTETEAVMGLPVSILTPEVVGCBELTGSNPFVTSID
VLIGITKHLRQGVAGKFVEFFGSGVSLSIDVDRITIANMCPYGAIIISFFPVDNVTI
KHLHETGFSKALSMETYLKAVKLFNRDQNSGDBPEXSOVIOINLSITVPSVSGPKR
PRDRAVTDMSKDFQACLNEKVFQIAAEKQKDIYSIHVEGSEYKLSHGSVVIAA
VLSCTNNCPNMYLAAGLLAKKAVEAGLRVKPYIRTSLSPGSGMTHYLLSSGVLPYL
KSLGFLPVGCGSTCGVNTAPLSDAVLNKVGKDLVTCGNFIWKNFEGRLCDCVRAN
YLASPLVYAYIAGVNTIDFQTEPLDGTQKNYLDHIDWPSREVRHEEVEHVLIS
MFKALKDKLEMGKNNSLEAPDSVLFPMDLKSTYIRCPSPFKLTKPFIALQATENA
HVLVLLKFNFTDHTSPAGSIARNSAAKLYLNRGLTPREFNSYCARGNDAVMTRGT
FANIKLFNFRLOPKPAPKTIHFPSGOTLDVFEAAELYQKEGIPLIILACKKYGSGNSRD
WAAKGYPYLVKAVLAESYEKIKHDKHLIGITAPLQFLPGENADSLGSLGRETFSLT
PEELSPGILNTQTSNGKVFVSIAEFDDVEITLYKHGGLNLFVARKFS"
BASE COUNT 1009 a 594 c 713 g 964 t
ORIGIN

Query Match 23.6%; Score 1023; DB 6; Length 3280;
Best Local Similarity 100.0%; Pred. No. 6.8e-205;
Matches 1023; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGACGCCCCAAAGCAGATACGCCTTTGAGTACTCTTATTGAAACATTAATGACAGT 60
|||||
Db 1 ATGGACGCCCCAAAGCAGATACGCCTTTGAGTACTCTTATTGAAACATTAATGACAGT 60
|||||
QY 61 TCACATAAGAGTCTTCGATGTATCTAAACTTGGCCACCAAGTATGATGTTCTGCCTTAC 120
|||||

RESULT 5
AX377520
LOCUS AX377520
DEFINITION Sequence 17 from Patent WO0212284.
ACCESSION AX377520
VERSION AX377520.1 GI:19573708

2867 bp DNA linear PAT 18-MAR-2002

KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 Kirsch, W.M., Lennart, A., Kellin, W.J., Kang, D.K., Levine, R.L. and Roualt, T.A.
TITLE Iron regulating protein-2 (irp-2) as a diagnostic for neurodegenerative disease
JOURNAL Patent: WO 021284-A 17 14-FEB-2002; Loma Linda University Medical Center (US) ; THE SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES (US)
FEATURES Location/Qualifiers
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 871 a 527 c 630 g 839 t
ORIGIN
Query Match 22.8%; Score 990; DB 6; Length 2867;
Best Local Similarity 100.08; Pred. No. 5.9e-198; Mismatches 0; Indels 0; Gaps 0;
Matches 990; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 34 TACCTATTGAACATTAATGACAGTTCACATAAGAAAGTTCCTCGATGTATCAAACTT 93
DB 1 TACCTATTGAACATTAATGACAGTTCACATAAGAAAGTTCCTCGATGTATCAAACTT 60
QY 94 GGCACCAAGTATGATGTTCTGCTTACTCAATACCGGCTTGTGGAGCTGCTGTAGCA 153
DB 61 GGCACCAAGTATGATGTTCTGCTTACTCAATACCGGCTTGTGGAGCTGCTGTAGCA 120
QY 154 AATTGTGATGGCTTTTAATGAAGAAGAAAGATGTTATGAACATTTTAGACTGGAACACC 213
DB 121 AATTGTGATGGCTTTTAATGAAGAAGAAAGATGTTATGAACATTTTAGACTGGAACACC 180
QY 214 AAACAAAGCAATGTTGAAGTGCCTTTTTCCTGCCGCTGTTCTTCTTCAAGATTTTACT 273
DB 181 AAACAAAGCAATGTTGAAGTGCCTTTTTCCTGCCGCTGTTCTTCTTCAAGATTTTACT 240
QY 274 GGAATACCAAGCAATGTTGAAGTGCCTTTTTCCTGCCGCTGTTCTTCTTCAAGATTTTACT 333
DB 241 GGAATACCAAGCAATGTTGAAGTGCCTTTTTCCTGCCGCTGTTCTTCTTCAAGATTTTACT 300
QY 334 GATCCTTGAGAAAGTCCATCCTGCTGTCGACAGATCTTTACAGTTGACCATCTTTTACAA 393
DB 301 GATCCTTGAGAAAGTCCATCCTGCTGTCGACAGATCTTTACAGTTGACCATCTTTTACAA 360
QY 394 ATTGACTTCAGTAAATGTGCAATACAGAATGCACCAATCTCTGGAGGTGGTGACCTGCAG 453
DB 361 ATTGACTTCAGTAAATGTGCAATACAGAATGCACCAATCTCTGGAGGTGGTGACCTGCAG 420
QY 454 AAGCAGGAAGCTCTCCACTTAAAGTGCAGCCTTAAGAACTTCCCTGCAGAGGCCAG 513
DB 421 AAGCAGGAAGCTCTCCACTTAAAGTGCAGCCTTAAGAACTTCCCTGCAGAGGCCAG 480
QY 514 ACTACCTGCCGAGGATCTTGTGATTCTGGAGAACTAGGCCGAACTCAGGAACATTTTCT 573
DB 481 ACTACCTGCCGAGGATCTTGTGATTCTGGAGAACTAGGCCGAACTCAGGAACATTTTCT 540
QY 574 TCGCAGATTGAGAAATACACCCATCTCTGTCCTTTTTCATTTGCACACAGTGCCTGAACCT 633
DB 541 TCGCAGATTGAGAAATACACCCATCTCTGTCCTTTTTCATTTGCACACAGTGCCTGAACCT 600
QY 634 GAACAGCTGTTAAAAATCAAGAAGTAAAGTTCGCGAGAAATCGAGAGAGGCTTCAGTTT 693
DB 601 GAACAGCTGTTAAAAATCAAGAAGTAAAGTTCGCGAGAAATCGAGAGAGGCTTCAGTTT 660
QY 694 TTTTAAGTGGAGTTTCAAGAGTTTAAAGAATGTGGCAGTATCCCTCCTGGAATGGAATG 753
DB 661 TTTTAAGTGGAGTTTCAAGAGTTTAAAGAATGTGGCAGTATCCCTCCTGGAATGGAATG 720
QY 754 GCTCATCAATAAATAGATTAGATTATTTGTCAAGAGTGGTTTTTGAAGAAAAACACCTCCTC 813

Db 721 GCTCATCAATAAAGTATTTGTCAAGAGTGGTTTTTGAAGAAAGACCTTCCTC 780
QY 814 TTCCAGACAGTCTAGTCGGCAGAGATTTCACACATAACGATGGTGAATGGTTTAGGGATT 873
Db 781 TTCCAGACAGTCTAGTCGGCAGAGATTTCACACATAACGATGGTGAATGGTTTAGGGATT 840
QY 874 CTGGGTTGGGGGTTGGAGCATTGAAACAGAACAGCAGTTATGCTTGGTCTGCCAGTTTCT 933
Db 841 CTGGGTTGGGGGTTGGAGCATTGAAACAGAACAGCAGTTATGCTTGGTCTGCCAGTTTCT 900
QY 934 CTTACTTTACAGAGTGGTGGATGTGAGTTAACTGGGTGATCAACACCTTTTGTGTACA 993
Db 901 CTTACTTTACAGAGTGGTGGATGTGAGTTAACTGGGTGATCAACACCTTTTGTGTACA 960
QY 994 TCCATAGATGTTGTTCTTGGTATTACAAAG 1023
Db 961 TCCATAGATGTTGTTCTTGGTATTACAAAG 990
RESULT 6
HSU20180 2867 bp mRNA linear PRI 26-OCT-1995
LOCUS Human iron-regulatory protein 2 (IRP2) mRNA, partial cds.
DEFINITION
ACCESSION U20180
VERSION U20180.1 GI:897580
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2867)
AUTHORS Guo, B., Brown, F.M., Phillips, J.D., Yu, Y. and Leibold, E.A.
TITLE Characterization and expression of iron regulatory protein 2 (IRP2). Presence of multiple IRP2 transcripts regulated by intracellular iron levels
JOURNAL J. Biol. Chem. 270 (28), 16529-16535 (1995)
MEDLINE 95348066
PUBMED 7622457
REFERENCE 2 (bases 1 to 2867)
AUTHORS Brown, F.M.
TITLE Direct Submission
JOURNAL Submitted (23-JAN-1995) Fritz M. Brown, Program in Hum. Mol. Biol. and Genetics, University of Utah, 3250 Bldg 533, Salt Lake City, UT 84112, USA
FEATURES Location/Qualifiers
source
1. .2867
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="PSLFB4a, PSLFB15c"
/sex="female"
/tissue_type="brain"
/dev_stage="fetus"
/notes="Similar to GenBank Accession No. U20181"
1. .2867
/gene="IRP2"
/gene="IRP2"
/gene="IRP2"
/codon_start=1
/product="iron-regulatory protein 2"
/protein_id="AAA79926.1"
/db_xref="GI:897581"
gene
CDS
/translation="YLITLNDSSHKFFDVKLGTGYDVLPSYIRVILEAAVRNCDG FLMKKEDVMNLDWTKQSNVEYFPFARVLLQDFTGIPAMVDFAAMREAVKTLGGDP EKYPACPTDLTDVHSLQIDFSKCAIQANAPNPGGDLQAGKLSPLKVPQPKLPGRGQ TTRCGSDSGELGRNSTSSQENTPILCPHLPQVPEPTVINKQEVFGNRRRL QEFKWSRVLKNVAVIPPGTGMARHINLEYSRVFEKDLFPDVSQVCTDSHTWVN GLGILWGVGGIETEAAMLGLPSLTPVVCCELTGSSNPFTSDDVLGITKHURQ VVAGKPEFFGSGVSQLSIVDRTITANMCPPEYGAILLSFPVDNVTLKHLEHTGFSKA KLESMETYLKAVLFRNDQSGEPEYSQVQINLSIVPSVSGPRPRDRVAVTDMK SDFQACINERKVGKFGQIAEKQKQDVSITHYEGSEYKLSHGVSIVAAVISCNNCRPS VMLAAGLLAKKAVEAGLRVKPYIRTSISPGSGMVTYLSSSGVLPSLKLGIYVYG CSTCVGNTAPLSDAVLNAVKGDLVTCGNFINKKNFEGRCLDCVRANYLASPPLVVAY

ALAGTWNIDFOTPLGTDPKNIYLIHDIWPSREEVHVEEHLVLSMFKALKOKIEM
GNKRNSLEAPDSVLPWDLKSTYIRCPSPFDKLTKEPIALQAIENAGVLLVLDGDSVT
TDRISPAGSIARNSAAKYLINRGLTPREFNSYGARRNDAYMTRGTFRANIKLNKFI
GKPAKRTIHFPSSQTLDFVEAAELIKGEGIPILIIILAGKKYSGNSRDWAAGKPYLLGV
KAVLAESYKIHKDLHIGTIGTAPLQPLFQENADSLGSLGRETFSLTFPEELSPGITLN
IOTSTGKFVSFTASFEDDDVEITLYRHGGLLNFAVKES"
BASE COUNT 871 a 527 c 630 g 839 t
ORIGIN

Query Match 22.8%; Score 990; DB 9; Length 2867;
Best Local Similarity 100.0%; Pred. No. 5.9e-198;
Matches 990; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 34 TACCTTTATCAAAACATTAATGACAGTTCACATAAGAAGTCTTCGATGTATCTAAACTT 93
Db 1 TACCTTTATGAACATTAATGACAGTTCACATAAGAAGTCTTCGATGTATCTAAACTT 60
Qy 94 GGCACCAAGTATGATGTTCTGCCTTACTCAATACGGGTCTTGTGGAAGTGTGTACGA 153
Db 61 GGCACCAAGTATGATGTTCTGCCTTACTCAATACGGGTCTTGTGGAAGTGTGTACGA 120
Qy 154 AATTGTGATGGCTTTTAAATGACAGAGGAAGATGTTATGAACATTTTACAGCTGGAAC 213
Db 121 AATTGTGATGGCTTTTAAATGACAGAGGAAGATGTTATGAACATTTTACAGCTGGAAC 180
Qy 214 AAACAAAGCAATGTTGAAGTGCCTTTTCCCTGCCCGTCTTCTTCAAGATTTTACT 273
Db 181 AAACAAAGCAATGTTGAAGTGCCTTTTCCCTGCCCGTCTTCTTCAAGATTTTACT 240
Qy 274 GGAATACCAAGCATGTTGGATTTTGTCTGCTATGAGGAGGAGCTGAAACCTCTTGGAGGT 333
Db 241 GGAATACCAAGCATGTTGGATTTTGTCTGCTATGAGGAGGAGCTGAAACCTCTTGGAGGT 300
Qy 334 .GATCCTGAGAAGTCCATCCTGCTGTGCCAGACATCTTACAGTTGACCATTTCTTTACAA 393
Db 301 GATCCTGAGAAGTCCATCCTGCTGTGCCAGACATCTTACAGTTGACCATTTCTTTACAA 360
Qy 394 ATTCACTTCAGTAAATGTGCAATACAGAATGCACCAATCCTGGAGTGTGTGACCTGCAG 453
Db 361 ATTCACTTCAGTAAATGTGCAATACAGAATGCACCAATCCTGGAGTGTGTGACCTGCAG 420
Qy 454 AAACAGAGAAAGCTCTCTCCACTTAAAGTGCAGCCTTAAGAAGCTTCCTCGACAGGCCAG 513
Db 421 AAACAGAGAAAGCTCTCTCCACTTAAAGTGCAGCCTTAAGAAGCTTCCTCGACAGGCCAG 480
Qy 514 ACTACCTGCCAGATCTTGTGATCTGGAGACTAGCCGGAACCTCAGGAACATTTTCT 573
Db 481 ACTACCTGCCAGATCTTGTGATCTGGAGACTAGCCGGAACCTCAGGAACATTTTCT 540
Qy 574 TCGCAGATTGAGATAACACCATCCTGTGCTTTTCATTTGCAACCAAGTGCCTGAACCT 633
Db 541 TCGCAGATTGAGATAACACCATCCTGTGCTTTTCATTTGCAACCAAGTGCCTGAACCT 600
Qy 634 GAAACAGTGTATAAATAATCAAGAAGTAGAATTCGGCAGAAATCGAGAGAGCTTCAGTTT 693
Db 601 GAAACAGTGTATAAATAATCAAGAAGTAGAATTCGGCAGAAATCGAGAGAGCTTCAGTTT 660
Qy 694 TTTAAGTGGAGTTCAAGAGTTTTAAAGAAATGTGCAGTGATCCCTCCTGCACTGGAGATG 753
Db 661 TTTAAGTGGAGTTCAAGAGTTTTAAAGAAATGTGCAGTGATCCCTCCTGCACTGGAGATG 720
Qy 754 GCTCATCAATATAACTTAGAATATTGTGCAAGAGTGGTTTTGAAGAAAAAGACCTCCTC 813
Db 721 GCTCATCAATATAACTTAGAATATTGTGCAAGAGTGGTTTTGAAGAAAAAGACCTCCTC 780
Qy 814 TTCCACAGACGTAGTGGCGACAGATTCACATAACGATGGTGAATGGTTTAGGGATT 873
Db 781 TTCCACAGACGTAGTGGCGACAGATTCACATAACGATGGTGAATGGTTTAGGGATT 840
Qy 874 CTGGGGTGGGGGTTGGAGGCATTGAACACAGACAGTATGCTTGGTCTGCCAGTTTCT 933
Db 841 CTGGGGTGGGGGTTGGAGGCATTGAACACAGACAGTATGCTTGGTCTGCCAGTTTCT 900

Qy 934 CTTACTTTACCAGAGGTGCTTGGATGTGAGTTAACTGGGTGCATCAAAACCCCTTTGTACA 993
Db 901 CTTACTTTACCAGAGGTGCTTGGATGTGAGTTAACTGGGTGCATCAAAACCCCTTTGTACA 960
Qy 994 TCCATAGATGTTGTTCTTGGTATTACAAAG 1023
Db 961 TCCATAGATGTTGTTCTTGGTATTACAAAG 990
RESULT 7
AK027033 3341 bp mRNA linear PRI 29-SEP-2000
LOCUS Homo sapiens CDNA: FLJ23380 fis, clone HEP16267, highly similar to
DEFINITION HUMIREBPA Human iron-responsive element-binding protein/iron
regulatory protein 2 (IRE-BP2/IRP2) mRNA.
ACCESSION AK027033
VERSION AK027033.1 GI:10440047
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens hepatoma cell_line:HepG2 cDNA to mRNA, clone_lib:HEP
clone:HEP16267.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (sites)
Kawabata,A., Hikiiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 3341)
Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
Direct Submission
Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure Analysis, Human
Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
Japan (E-mail:cdnal@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing; Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing; Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).
FEATURES
source Location/Qualifiers
1..3341
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HEP16267"
/cell_line="HepG2"
/cell_type="hepatoma"
/clone_lib="HEP"
/note="cloning vector pME18SFL3"
misc_feature 1..3341
/note="highly similar to HUMIREBPA Human iron-responsive
element-binding protein/iron regulatory protein 2
(IRE-BP2/IRP2) mRNA"
BASE COUNT 1031 a 597 c 725 g 988 t
ORIGIN
Query Match 21.7%; Score 944.2; DB 9; Length 3341;
Best Local Similarity 99.7%; Pred. No. 2.6e-188;
Matches 946; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 75 CTTTCGATGATCTAAACTTGGCACCAAGATGATGTTCTGCCTTACTCAATACGGGTCTT 134
Db 1 CTTTCGATGATCTAAACTTGGCACCAAGATGATGTTCTGCCTTACTCAATACGGGTCTT 60
Qy 135 GTTGGAGCTGCTACGAAATTTGATGCTTTTAAATGAAGAAGAGATGTTATGAA 194
Db 61 GTTGGAGCTGCTGACGAAATTTGATGCTTTTAAATGAAGAAGAGATGTTATGAA 120

QY	195	CATTTTAGACTGAAACCAACAAAGCAATGTTGAAGTGCCCTTTTCCCTGCCCGTGT	254
Db	121	CATTTTAGACTGAAACCAACAAAGCAATGTTGAAGTGCCCTTTTCCCTGCCCGTGT	180
QY	255	TCCTCTTCAAGATTTTACTGGAATACACAAATGGTGGATTTTGCTGCTATCAGGAGGC	314
Db	181	TCCTCTTCAAGATTTTACTGGAATACACAAATGGTGGATTTTGCTGCTATCAGGAGGC	240
QY	315	AGTGAACCTCTTGGAGGTGATCCTCGAAGAGTCCATCCTGCTGTGTCGACAGATCTTAC	374
Db	241	AGTGAACCTCTTGGAGGTGATCCTCGAAGAGTCCATCCTGCTGTGTCGACAGATCTTAC	300
QY	375	AGTGAACCTCTTGGAGGTGATCCTCGAAGAGTCCATCCTGCTGTGTCGACAGATCTTAC	434
Db	301	AGTGAACCTCTTGGAGGTGATCCTCGAAGAGTCCATCCTGCTGTGTCGACAGATCTTAC	360
QY	435	TGGAGGTGGTGCCTGCAGAAAGCAGAAAGCTCTCTCCACCTTAAAGTGCAGCCTAAGAA	494
Db	361	TGGAGGTGGTGCCTGCAGAAAGCAGAAAGCTCTCTCCACCTTAAAGTGCAGCCTAAGAA	420
QY	495	GCTTCCCTCGAGAGGCCAGACTACCTGCCGAGGATCTTGTGATCTGGAGAACTAGGCGG	554
Db	421	GCTTCCCTCGAGAGGCCAGACTACCTGCCGAGGATCTTGTGATCTGGAGAACTAGGCGG	480
QY	555	AAACTCAGAACATTTTCTTCCAGATTCAGAAATACACCATCCTGTCCTTTTCATTT	614
Db	481	AAACTCAGAACATTTTCTTCCAGATTCAGAAATACACCATCCTGTCCTTTTCATTT	540
QY	615	GCACCAAGTGCCTGAAACCTGAAACAGTGTAAATAATCAAGAGTAGAATTCGGCAGAAA	674
Db	541	GCACCAAGTGCCTGAAACCTGAAACAGTGTAAATAATCAAGAGTAGAATTCGGCAGAAA	600
QY	675	TCGAGAGAGGCTTCAGTTTTTAAAGTGGAGTTCAAGAGTTTTTAAAGATTCGGCAGTAT	734
Db	601	TCGAGAGAGGCTTCAGTTTTTAAAGTGGAGTTCAAGAGTTTTTAAAGATTCGGCAGTAT	660
QY	735	CCCTCTCGAACTGGAATGGCTATCAATAAATTTAGATTTTGTCAAGAGTGGTTTT	794
Db	661	CCCTCTCGAACTGGAATGGCTATCAATAAATTTAGATTTTGTCAAGAGTGGTTTT	720
QY	795	TGAAGAAAAGACCTCCTCTCCAGACAGTGTAGTCCGACAGATTCACACATAACGAT	854
Db	721	TGAAGAAAAGACCTCCTCTCCAGACAGTGTAGTCCGACAGATTCACACATAACGAT	780
QY	855	GGTGAATGTTTTAGGATTTCTGGGTGGGGGTTGGAGGCATTTGAACAGAGCAGTTAT	914
Db	781	GGTGAATGTTTTAGGATTTCTGGGTGGGGGTTGGAGGCATTTGAACAGAGCAGTTAT	840
QY	915	GCTTGGTCTGCCAGTTTCTCTTACTTTACCAGAGTGGTTGGATGTAGTTAACTGGGTC	974
Db	841	GCTTGGTCTGCCAGTTTCTCTTACTTTACCAGAGTGGTTGGATGTAGTTAACTGGGTC	900
QY	975	ATCAAAACCTTTTGTATACATCATAGATGTTGTTCTTGGTATTACAAAG 1023	
Db	901	ATCAAAACCTTTTGTATACATCATAGATGTTGTTCTTGGTATTACAAAG 949	
RESULT	8		
AX377522			
LOCUS	AX377522	3770 bp	DNA linear PAT 18-MAR-2002
DEFINITION	Sequence 19 from Patent WO0212284.		
ACCESSION	AX377522		
VERSION	AX377522.1	GI:19573709	
KEYWORDS			
SOURCE	Norway rat.		
ORGANISM	Rattus norvegicus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;		
	Rattus.		
	1		
	Kirsch,W.M., Lennart,A., Kelln,W.J., Kang,D.K., Levine,R.L. and		
	Roualt,T.A.		
TITLE	Iron regulating protein-2 (lrp-2) as a diagnostic for		

JOURNAL	neurodegenerative disease		
FEATURES	Patent: WO 0212284-A 19 14-FEB-2002;		
source	Loma Linda University Medical Center (US) : THE SECRETARY OF THE		
	DEPARTMENT OF HEALTH AND HUMAN SERVICES (US)		
	Location/Qualifiers		
	1. .3770		
	/organism="Rattus norvegicus"		
	/db_xref="taxon:10116"		
BASE COUNT	1119 a 713 c 840 g 1098 t		
ORIGIN			
Query Match	20.2%; Score 877.2; DB 6; Length 3770;		
Best Local Similarity	91.6%; Pred. No. 3.2e-174;		
Matches	940; Conservative	0; Mismatches	83; Indels
	3; Gaps		1;
QY	1	ATGACGCCCCCAAGACGAGGATACGCCCTTTGAGTACCTTATTGAAACATTTAAATGACAGT	60
Db	12	ATGGACTCCCCCAAGTGGAGGATACACCTTTGAGTACCTTATTGAAACATTTAAATGACAGT	71
QY	61	TCACATAGAAGTTCTTTCGATGTATCTAACTT---GGCACCAAGATATGATGTTCTGSCCT	117
Db	72	TCACAGAAGAAGTTCTTCAATGTACCTTAACTTGGAGGACCAAGATATGATATCTGSCCT	131
QY	118	TACTCAATACGGTCTTGTGGAAAGCTGTGTACGAAATTTGTATGCTGTTTAAATGAAG	177
Db	132	TACTCAATACGGTCTTGTGGAAAGCTGTGTACGAAATTTGTATGCTGTTTAAATGAAG	191
QY	178	AAGAAAGATGTTATGAACATTTTAGACTGGAACACCAACAAAGCAATGTTGAAGTGCCC	237
Db	192	AAGAAAGATGTTATGAACATTTTAGACTGGAACACCAACAAAGCAATGTTGAAGTGCCC	251
QY	238	TTTTTCCCTGCCCGTCTTCTTCAAGATTTTACTGGAATACAGCAATGTTGGATTTT	297
Db	252	TTTTTCCCGCCCGTGTCTTCAAGATTTTACTGGAATACCGCAATGTTGGATTTT	311
QY	298	GCTGCTATGAGGGAGGAGTGAACACTCTTGGAGGTGATCCTCAGAAAGTCCCTCCTGCT	357
Db	312	GCTGCTATGAGGGAGGAGTGAACACTCTTGGAGGTGATCCTCAGAAAGTCCCTCCTGCT	371
QY	358	TGTCGACACATCTTACAGTTGACCACTCTTTACAGATTTGACCTTAAATGTCGAATA	417
Db	372	TGTCGACACATCTTACAGTTGACCACTCTTTACAGATTTGACCTTAAATGTCGAATA	431
QY	418	CAGAATCACCAATCCTCGAGGTGTTGACCTGCAGAAAGCAGGAAGTCTCTCCACTT	477
Db	432	CAGAATCACCAATCCTCGAGGTGTTGACCTGCAGAAAGCAGGAAGTCTCTCCACTT	491
QY	478	AAAGTGAGCCTTAAGAAGCTTCCCTCGAGGCGCAGACTACCTGCGGAGAGTCTTGTGAT	537
Db	492	AAAGTGAGCCTTAAGAAGCTTCCCTCGAGGCGCAGACTACCTGCGGAGAGTCTTGTGAT	551
QY	538	TCGGAAGACTAGGCGGAAACTCAGGAACATTTTCTTCGAGATTCAGAAATACACCATC	597
Db	552	TCGGAAGACTAGGCGGAAACTCAGGAACATTTTCTTCGAGATTCAGAAATACACCATC	611
QY	598	CTGTGCTCTTTTCATTTGCAACCAAGTGCCTGAACTTGAACAGTGTAAAAAATCAAGAA	657
Db	612	CTGTGCTCTTTTCATTTGCAACCAAGTGCCTGAACTTGAACAGTGTAAAAAATCAAGAA	671
QY	658	GTAGAATTCGGCAGAAATCGAGAGAGCTTTCAGTTTAAAGTGGAGTTCAAGAGTTTAA	717
Db	672	GTAGAATTCGGCAGAAATCGAGAGAGCTTTCAGTTTAAAGTGGAGTTCAAGAGTTTAA	731
QY	718	AAGAATGTGGCAGTGTCTCTCTCGAAGTGGAAATGGCTCATCAAAATTAACCTTAGAATAT	777
Db	732	AAGAATGTGGCAGTGTCTCTCTCGAAGTGGAAATGGCTCATCAAAATTAACCTTAGAATAT	791
QY	778	TTGTCAGAGTGTGTTTTGAAGAAAAGACCTCCTCTCCAGACAGTGTAGTCGACACA	837
Db	792	TTGTCAGAGTGTGTTTTGAAGAAAAGACCTCCTCTCCAGACAGTGTAGTCGACACA	851
QY	838	GATTACACATACGATGTTGATGTTTGGATTTCTGGGTGGGGGTTGGAGGCATT	897

Db 852 GATTCTCATATAACCATGGTGAATGGGAATCTTGGTGGGAGTTGGAGGCATT 911

QY 898 GAAACAGAGCAGTATCTGGTGGTCCAGTCTCTTACTTACATTTACCAGAGGTGGTGGGA 957

Db 912 GAGACAGAGGAGTATCTGGTGGTCCAGTCTTACTTACTTACCAGAGGTGGTGGGA 971

QY 958 TGTGAGTTAACTGAGTCACTAAACCTTTTGTATACATCCATAGATGTTGTTCTTGGTATT 1017

Db 972 TGTGAGTAACTGGTCACTCAATGCTTTGTTACATCCATAGATGTTGTTCTTAGGCATT 1031

QY 1018 ACAAG 1023

Db 1032 ACAAG 1037

RESULT 9

RNU20181

LOCUS Rattus norvegicus iron-regulatory protein 2 (IRP2), mRNA, complete cds.

DEFINITION RNU20181 3770 bp mRNA linear ROD 26-OCT-1995

ACCESSION U20181.1 GI:897582

VERSION U20181.1

KEYWORDS

SOURCE rat.

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

1 (bases 1 to 3770)

Guo,B., Brown,F.M., Phillips,J.D., Yu,X. and Leibold,E.A.

Characterization and expression of iron regulatory protein 2 (IRP2). Presence of multiple IRP2 transcripts regulated by intracellular iron levels

J. Biol. Chem. 270 (28), 16529-16535 (1995)

95348066

7622457

REFERENCE 2 (bases 1 to 3770)

Brown,F.M.

Direct Submission

Submitted (23-JAN-1995) Fritz M. Brown, Program in Hum. Mol. Biol. and Genetics, University of Utah, 3250 Bldg 533, Salt Lake City, UT 84112, USA

FEATURES

location/Qualifiers

1..3770

/organism="Rattus norvegicus"

/strain="Sprague-Dawley"

/db_xref="taxon:10116"

/clone="pSL800, pU2B2-1"

/sex="male"

/tissue_type="liver"

/dev_stage="adult"

/note="Similar to GenBank Accession No. U20180"

1..3770

/gene="IRP2"

12..2903

/gene="IRP2"

/codon_start=1

/product="iron-regulatory protein 2"

/protein_id="AA79927.1"

/db_xref="GI:897583"

/translation="MDSPSAGTYFEYLIEITNLGSSQKFFNVPKLGGTKYDILPYRIR VLEAAVTRNGDGLMKEDVINLDWTKQSNVEVFPFPAVVLDQFTGIPAMVDFAA MRAMATLGGDKPKVHPACPTDLTVDSHSLQIDFSKAIQANPNPBGDLDQKAGKLSPL KVPQKPLRGQDTCRGSQDSGELNSGTFSSOIENTVLCFPHLPQVPEPETVLKN QVEFGNRRERLOFQKWSGAPKFNVAIPPGTCMAHVNLEHLSRVVFEADLLFPDS VIGTDSHITWNLGILGWGVGIEEAVMLGLPVLTLPEVVGCELTGSSNAFTVSI DIVLGHITKHLRGVAGKEVFPFGVSQSLIVDRFTIANMCPCEYGAILSFPPVONT LRLEHVTGDKTLESMEYLKAKLFRNDENSEPEYSQVIQINLNSIVASVGPFR PODRVATDMKSDFOACLNEKVGFGQVAEEKQSDTVSVRYDGSYKLSHGSVIAA VISCTNNCPNMLAAGLLAKAVETGLRVKPYIRTSLSGSGMVTXYLSSSVLPYL SKLGFELVGVGCTCVGNTPALSEAILNNAVKQGDLCATCVLSGNKNFGRLCDVRAN YLASFPLVAYATAGTVNIDFQTEPLGCTDSTGKNLYLHDHIFPSREHVEHILS MFALAKEVMGNKRNWNSLDAPDSVLFPPWDPVKSTYIRCPSPFFDKLTKEPAASQPIENA

HVLLYLGDVSTTDHISPAGSIARSAARAAKYLTNRGLTPREFNSYGARNGNDVWTRGT FANIKFNFKIPKPAKPTIHFPGOTLDVFEAAELYQKEGIPLIILAKKKYSGNSRD WAAGKPYLLGVKRAVLAESYERIKHDLHIGIAPLEFLFPGENADSLGSLGREVFSLSF PBELPFGITLNIKTSTGKEFSVIAAFENDVEITLYKHGGLLNFVARKFL"

BASE COUNT 1119 a 713 c 840 g 1098 t

ORIGIN

Query Match 20.2%; Score 877.2; DB 10; Length 3770;

Best Local Similarity 91.6%; Pred. No. 3.2e-174;

Matches 940; Conservative 0; Mismatches 83; Indels 3; Gaps 1;

QY 1 ATGGACGCCCAAAAGCAGGATACGCTTTGAGTACCTTATTCACAACTTAATGACAGT 60

Db 12 ATGGACTCCCAAGTCAGGATACACCTTTGAGTACCTTATTCACAACTTAATGACAGT 71

QY 61 TCACATAAGAAAGTCTTCGATGATCTAACTT---GGCACCAAGTATGATGTTCTGCCT 117

Db 72 TCACAGAAGAAGTCTTCAATGTACCTAACTTTGGAGGCCACCAAGTATGATATCTGCCT 131

QY 118 TACTCAATACGGCTCTTGTGGAGCTGCTGTACGAAATTTGATGGCTTTTAAATGAAG 177

Db 132 TACTCAATACGGCTCTTGTGGAGCTGCTGTACGAAATTTGATGGCTTTTAAATGAAG 191

QY 178 AAGGAAGATGTTAAGCAATTTTAGACTGGAAACCAAAACCAAGCAATTTGAAGTGCCC 237

Db 192 AAGGAAGATGTTAATAATTTTGGACTGGAAACCAAAACCAAGCAATTTGAAGTGCCC 251

QY 238 TTTTCCCTCCCGTGTCTTCTCAAGATTTTACGGAATACCAAGTATGTTGGATTTT 297

Db 252 TTTTCCCGCTGTTGTTCTTCAAGATTTTCACTTGAATACCGCAATTTGGATTTT 311

QY 298 GCTGCTATGAGGAGGACGAGTGAAGTCTTGGAGGTGATCCCTCAGAAAGTCCATCTGCT 357

Db 312 GCTGCTATGAGGAGGACGAGTGAAGTCTTGGAGGTGATCCCTAAGAAAGTCCACCTGCC 371

QY 358 TGTCCGACAGATCTTACAGTTGACCATTTCTTACAAATTTGACTTCAAGTAAATGTGAATA 417

Db 372 TGTCCAAACAGATCTCAGATTGACCATCTTTACAGATTGACTTCAAGTAAATGTGAATA 431

QY 418 CAGAAATGCACCAATCTTGGAGTGTGACCTGCAGAAAGCAAGGAAAGTCTCTCCACTT 477

Db 432 CAGAAATGCACCAATCTTGGAGTGTGACCTGCAGAAAGCAAGGAAAGTCTCTCCACTT 491

QY 478 AAAGTGCAGCTAAGAAAGTCTTCCCTGCAGAGGCCACACTACCTGCCGAGAGTCTTGTGAT 537

Db 492 AAAGTGCAGCTAAGAAAGTCTTCCATGTGAGGCCACACTACCTGCCGAGAGTCTTGTGAT 551

QY 538 TCTGGAGAACTAGGCCAAACTCAGGAACATTTTCTTCGAGATTGAGAAATACACCATC 597

Db 552 TCTGGAGAACTAAGCCGAAACTCAGGAACATTTTCTTCGAGATTGAGAAATACACCATC 611

QY 598 CTGTGCTCTTTTCATTTGCAACAGTGCCTGAACCTGAAACAGTGTAAAAATCAAGAA 657

Db 612 CTGTGCTCTTTTCATTTGCAACAGTGCCTGAACCTGAAACAGTGTAAAAATCAAGAA 671

QY 658 GTAGAATTCGGCAGAAATCGAGAGGCTTTCAGTTTAAAGTGGAGTTCAGAGTCTTTA 717

Db 672 GTAGAATTCGGCAGAAATCGAGAGGCTTTCAGTTTAAAGTGGAGTTCAGAGTCTTT 731

QY 718 AAGAAATGGCAGTGAATTCCTTGGAACTGGAATGGCTCATCAATAAATAGAAAT 777

Db 732 AAGAAATGGCAGTGAATTCCTTGGAACTGGAATGGCTCATCAATAAATAGAAAT 791

QY 778 TTGTCAGAGTGGTTTTTCAAGAAAGACCTTCTCCACAGACAGTGTAGTCGCGACA 837

Db 792 TTGTCAGAGTGGTTTTTCAAGAAAGACCTTCTCCACAGACAGTGTAGTCGCGACA 851

QY 838 GATTTCACATTAACGATGCTGAATGTTTAGGATTTCTGGGTTGGGGTCTGGAGCAT 897

Db 852 GATTCTCATATACCATGCTGAATGGAATTTCTTGGGTTGGGAGTTCGGAGCAT 911

QY 898 GAAACAGAGCAGTATGCTTGGTCTGCCAGTCTTCTTACTTTACCAGAGTGTGTGA 957


```
Db 912 GAGACAGAGCGAGTATGCTTGGCCCTGCAGTTACTCTTACTTTTACCAGAGGTGGTTGGA 971
QY 958 TGTGAGCTTAACGGGTCATCAACCCCTTTTGTACATCATAGATGTTCTTCTTGGTATT 1017
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 972 TGTGAGCTTAACGGGTCATCCCACTGCTTTTGTATCATCCATAGATATTGTCCTAGGCAVT 1031
QY 1018 ACAAG 1023
      |||||
Db 1032 ACAAG 1037

RESULT 10
AC004469/c
LOCUS
DEFINITION
Homo sapiens clone CTB-58E15, WORKING DRAFT SEQUENCE, 126 unordered
pieces.
AC004469
AC004469.8 GI:11141978
HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE
Homo sapiens.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 309662)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbaria,J.,
Benton,J., Blinag,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Denn,A.C., Detherick,D.A., Delaney,K.R., Delgado,O.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Loulseg,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenwo,S.,
Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Picken,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,
Sisson,I., Sodergren,E., Sonaik,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Watlingon,S., Williams,G., Williams,A., Wleszyk,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.
and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 309662)
Worley,K.C.
Direct Submission
Submitted (21-MAR-1998) Molecular and Human Genetics, Baylor
College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 14, 2000 this sequence version replaced gi:11079386.
----- Genome Center
Center: Baylor College of Medicine
```

```
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: WO
Center clone name: CTB-58E15
----- Summary Statistics
Sequencing vector: M13; L08821
Chemistry: Dye-primer Bodipy: 10% of reads
Chemistry: Dye-terminator Big Dye: 24% of reads
Assembly program: Phrap: version 0.990329
Consensus quality: 263288 bases at least Q40
Consensus quality: 281803 bases at least Q30
Consensus quality: 292338 bases at least Q20
Estimated insert size: 261499; sum-of-ctigs estimation
Quality coverage: 12.9x in Q20 bases; sum-of-ctigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
* NOTE: This sequence may represent more than one clone.
* NOTE: This is a 'working draft' sequence. It currently
consists of 126 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
* 1 7683: contig of 7683 bp in length
* 7684 7783: gap of unknown length
* 7784 12646: contig of 4863 bp in length
* 12647 12746: gap of unknown length
* 12747 19521: contig of 6775 bp in length
* 19522 19621: gap of unknown length
* 19622 26379: contig of 6758 bp in length
* 26380 26479: gap of unknown length
* 26480 32211: contig of 5731 bp in length
* 32211 32311: gap of unknown length
* 32311 38553: contig of 6243 bp in length
* 38554 38653: gap of unknown length
* 38654 42723: contig of 4070 bp in length
* 42724 42823: gap of unknown length
* 42824 45577: contig of 2752 bp in length
* 45578 45676: gap of unknown length
* 45677 50160: contig of 4485 bp in length
* 50161 50260: gap of unknown length
* 50261 54295: contig of 4035 bp in length
* 54296 54396: gap of unknown length
* 54397 59662: contig of 5266 bp in length
* 59663 63562: gap of unknown length
* 63563 63662: contig of 3801 bp in length
* 63663 67444: gap of unknown length
* 67445 72937: contig of 3782 bp in length
* 72938 73038: gap of unknown length
* 73039 77391: contig of 4254 bp in length
* 77392 81075: gap of unknown length
* 81076 81176: contig of 3684 bp in length
* 81177 84197: gap of unknown length
* 84198 88352: contig of 3021 bp in length
* 88353 88452: gap of unknown length
* 88453 91923: contig of 4055 bp in length
* 91924 92023: contig of 3472 bp in length
* 92024 96830: gap of unknown length
* 96831 96931: contig of 4807 bp in length
* 96932 101888: gap of unknown length
* 101889 101989: contig of 4958 bp in length
* 101990 105610: gap of unknown length
* 105611 105611: contig of 3522 bp in length
* 105612 108586: gap of unknown length
* 108587 108586: contig of 2976 bp in length
```


* 108587	108586:	gap of unknown length	
* 108687	111359:	contig of 2673 bp in length	
* 111360	111459:	gap of unknown length	
* 111460	115189:	contig of 3730 bp in length	
* 115190	115289:	gap of unknown length	
* 115290	118441:	contig of 3152 bp in length	
* 118442	118541:	gap of unknown length	
* 118542	121774:	contig of 3233 bp in length	
* 121775	121874:	gap of unknown length	
* 121875	124980:	contig of 3106 bp in length	
* 124981	125080:	gap of unknown length	
* 125081	128388:	contig of 3308 bp in length	
* 128389	128488:	gap of unknown length	
* 128489	132514:	contig of 4026 bp in length	
* 132515	132614:	gap of unknown length	
* 132615	136097:	contig of 3483 bp in length	
* 136098	136197:	gap of unknown length	
* 136198	139733:	contig of 3536 bp in length	
* 139734	139833:	gap of unknown length	
* 139834	141892:	contig of 2059 bp in length	
* 141893	141992:	gap of unknown length	
* 141993	144800:	contig of 2808 bp in length	
* 144801	144900:	gap of unknown length	
* 144901	148527:	contig of 3627 bp in length	
* 148528	148627:	gap of unknown length	
* 148628	151818:	contig of 3191 bp in length	
* 151819	151918:	gap of unknown length	
* 151919	155407:	contig of 3489 bp in length	
* 155408	155507:	gap of unknown length	
* 155508	156847:	contig of 1140 bp in length	
* 156848	156747:	gap of unknown length	
* 156748	161063:	contig of 4316 bp in length	
* 161064	161163:	gap of unknown length	
* 161164	164289:	contig of 3126 bp in length	
* 164290	164389:	gap of unknown length	
* 164390	167923:	contig of 3534 bp in length	
* 167924	168023:	gap of unknown length	
* 168024	172057:	contig of 4034 bp in length	
* 172058	172157:	gap of unknown length	
* 172158	174285:	contig of 2128 bp in length	
* 174286	174385:	gap of unknown length	
* 174386	177135:	contig of 2750 bp in length	
* 177136	177235:	gap of unknown length	
* 177236	179247:	contig of 2012 bp in length	
* 179248	179347:	gap of unknown length	
* 179348	182490:	contig of 3143 bp in length	
* 182491	182590:	gap of unknown length	
* 182591	185132:	contig of 2542 bp in length	
* 185133	185232:	gap of unknown length	
* 185233	187603:	contig of 2371 bp in length	
* 187604	187703:	gap of unknown length	
* 187704	190087:	contig of 2384 bp in length	
* 190088	190187:	gap of unknown length	
* 190188	192719:	contig of 2532 bp in length	
* 192720	192819:	gap of unknown length	
* 192820	194744:	contig of 1925 bp in length	
* 194745	194844:	gap of unknown length	
* 194845	197287:	contig of 2443 bp in length	
* 197288	197387:	gap of unknown length	
* 197388	200098:	contig of 2711 bp in length	
* 200099	200198:	gap of unknown length	
* 200199	202373:	contig of 2175 bp in length	
* 202374	202473:	gap of unknown length	
* 202474	204607:	contig of 2134 bp in length	
* 204608	204707:	gap of unknown length	
* 204708	206234:	contig of 1527 bp in length	
* 206235	206334:	gap of unknown length	

Query Match 9.1%; Score 396.2; DB 2; Length 309662;
 Best Local Similarity 57.1%; Pred. No. 5.8e-73;
 Matches 701; Conservative 1; Mismatches 474; Indels 51; Gaps 3;
 QY 2360 TTGGTTTGAATATATATCTGTCTTCTTTTCATCTCTTAAATAAATAAACCATAGGCGCG 2419

Query Match 9.1%; Score 396.2; DB 2; Length 309862;
Best Local Similarity 57.1%; Pred. No. 5.8e-73;
Matches 701; Conservative 1; Mismatches 474; Indels 51; Gaps 3;
QY 2360 TTGGTGGTTGTAATATCTAATCTGTGTTCTTTTCATCTATAAAATAAACCATAGSCGG 2419

misc_feature	/db_xref="taxon:9606" /chromosome="4" /map="4"	
misc_feature	/clone="RP11-745L13" /clone_lib="RPC1-11 Human Male BAC" 1..1421	
misc_feature	/note="assembly_fragment" 1522..3070	
misc_feature	/note="assembly_fragment" 3171..4619	
misc_feature	/note="assembly_fragment" 4720..6092	
misc_feature	/note="assembly_fragment" 6193..7416	
misc_feature	/note="assembly_fragment" 7517..9502	
misc_feature	/note="assembly_fragment" 9603..11758	
misc_feature	/note="assembly_fragment" 11859..14447	
misc_feature	/note="assembly_fragment" 14548..16648	
misc_feature	/note="assembly_fragment" 16749..19136	
misc_feature	/note="assembly_fragment" 19237..22543	
misc_feature	/note="assembly_fragment" clone_end:SP6 vector_side:left	
misc_feature	/note="assembly_fragment" 22644..26220	
misc_feature	/note="assembly_fragment" 26321..30523	
misc_feature	/note="assembly_fragment" 30624..33581	
misc_feature	/note="assembly_fragment" 33682..38440	
misc_feature	/note="assembly_fragment" 38541..43644	
misc_feature	/note="assembly_fragment" 43745..49564	
misc_feature	/note="assembly_fragment" 49665..55504	
misc_feature	/note="assembly_fragment" 55605..60608	
misc_feature	/note="assembly_fragment" 60709..66316	
misc_feature	/note="assembly_fragment" 66417..73994	
misc_feature	/note="assembly_fragment" 74095..82238	
misc_feature	/note="assembly_fragment" 82339..86919	
misc_feature	/note="assembly_fragment" clone_end:T7 vector_side:left	
misc_feature	/note="assembly_fragment" 87020..93376	
misc_feature	/note="assembly_fragment" 93477..102272	
misc_feature	/note="assembly_fragment" 102373..112950	
misc_feature	/note="assembly_fragment" 113051..125539	
misc_feature	/note="assembly_fragment" 125640..138701	
misc_feature	/note="assembly_fragment" 138802..153801	
misc_feature	/note="assembly_fragment"	
Query Match	8.98; Score 385.4; DB 2; Length 176314;	
Best Local Similarity	62.59; Pred. No. 1.1e-70;	
Matches 702; Conservative	1; Mismatches 382; Indels 38; Gaps 5;	
Qy 2391 TCATCCTAAAAATAAAACCATAGCGCGGGTGGCTCACGCCCTGTATCCAGCAC	2450	

Db 12402	TCAGTCAAAAAATAAAGGCTGGAGCGCGGTGGCTCAGCGCTGTAAATCCAGCAC	12461
Qy 2451	TTTGGAGCGCGAGGCGGTGGATCATGAGTCAGGAGATCGAGACCATCTCTGGCTAACA	2510
Db 12462	TTTGGAGCGCGAGGCGGTGGATCATGAGTCAGGAAATCGAGACCATCTCTGGCTAACA	12521
Qy 2511	AGGTGAAACCCCGTCTCTACTAAAAATACAAAAATAGCCGGCGCGTGGCGGCGCC	2570
Db 12522	AGGTGAAACCCCATCTCTATTAATAACAAAAATAGCCGGCGCGTGGCGGCGCC	12581
Qy 2571	TGTAGTCCCGAGCTTCTCGGAGGCTGAGCGAGAGATGGCTGAACCCGGAGGCGGAG	2630
Db 12582	TGTAGTCCCGAGCTCTCGGAGGCTGAGCGAGAGATGGCTGAACCCGGAGGCGGAG	12641
Qy 2631	CTTGCAGTGAGCAAGATTGCGCCACTGCAGTCCCGAGTCCCGGCTGGCGGACAGCGA	2690
Db 12642	CTTGCAGTGAGCGGAGATTGCGCCACTGCAGTCCCGAGTCCCGGCTGGCGGACAGCGA	12701
Qy 2691	GACTCTGTCTCAAAAAAATAAAAAAATAAAATGAGGAAACGCATCTTT	2750
Db 12702	GACTCGCTCTCAAAAAATAAAATAAAAAAATAAAGGCTGGATGAACCCCTT	12761
Qy 2751	ACACTTAGGCTTTGAGTTTCTGTATCTATAAAAAAGGTTTGGATTAAGTATGATCCCTGCGC	2810
Db 12762	T-----TGAAAGCAGGGTTTTTACTCGGTTTTTTCATTGTT	12796
Qy 2811	ACTTATAAATGTTAGGCTTAATATTAATATATAGATCAGGATAGTTTCATCTTTAGTC	2870
Db 12797	ACATCCAGATGCTACACAGCATCTGAGACATTTTAGGTTTCATGATCATGATGCAATTTA	12856
Qy 2871	GCCTCCTTAGTCACCTCTCCATACCAATCTGAG--ACCATTTTCAATTTAGAAAAGAC	2928
Db 12857	GTGAATTGAGTTTATTTTCTTAGAATAAAGAGGTTAAGATTAATATACAAGTTAC	12916
Qy 2929	AAATAACTGTTGGTTACTTGTATAGTATAATAACCAAGAAAAATAATTTAGAGGAAAT	2988
Db 12917	AATGGGACTATGGAGCAATGAAGACCTTTTTCCTTTTTCCTTTTTCCTTTTTCCTTTT	12976
Qy 2989	TAAAGTTGAAACCATGTTTAAACAATTTACCAAGTGGGATTTGCCCTGTGATTAAAGA	3048
Db 12977	CTGGATATATCTTAAAAAATAAAATTCATAATGTTGAGGAAAAATAAATAATTTAAAG	13036
Qy 3049	TGCTGTAAACATTTGGCCAGTAGTTATTAATTTGAAAAATGTTTATAGCAATATATAT	3108
Db 13037	CTTTCAAAAAATAAATAACCATGTACCACAGAAAAATAAATAATGCAATGACCCAGAT	13096
Qy 3109	TTTTTATTTAAATATACAGTTTTCATCAGTCTATTAGTATTTCATTAAGTCTTAAGATGCCA	3168
Db 13097	ATTGTGCTCAGTGAATAAAATTCGCTCTAAGTATATATTAGCATTTTGTGAAACTGAT	13156
Qy 3169	TCAGTGGTTAGCAACACCATCTGTTTATGCTACTGCTAGAGAAGATAAAGGCTGTGTG	3228
Db 13157	CACATCGCAAGTGACAAAAATTTGCAATAAATTTGATAAAGTAGTGTCTCACACCTGT---	13213
Qy 3229	CAGTGGCTCACACCTGTGGAGCCCAAGCAGSAGCATCACCTTGAGCCAGAGTTCAAG	3288
Db 13214	---AATCCCATCACCTTTGGAGGCCAAGCGGGCAGATCACCTTGAGGTGAGGATTTAAA	13270
Qy 3289	ACCAACCTGCTCAACATTTGTAAGACCCCTCTCTACAAAAAATAAATAAATAAATAAATAA	3348
Db 13271	ACCAGCCTGCTCAACATGCTGAAGTCTTCTTCTACCAAAAAATA---CAAAATTAG	13326
Qy 3349	CTGGGTGCGGTGGACATGCCCTGTAGTCTCCAGTCTCTGAGAGGCTTAAGTGGGAGGATT	3408
Db 13327	CTGGGTGCTGCTGATGCCCTGTAGTCCAGCTAGTTGGCAGCTGATGTGGAGAAATT	13386
Qy 3409	GCTAGAGCCACGGTGTGGAAGCTGCAA--TGAGCTGTGACCACACACCTGCCCTCCAGCG	3467
Db 13387	GCTTGAACCTGGGAGCGAGAGGCTACAAGTAAGCTGAGAGCACACCTGCATCCAGCC	13446
Qy 3468	TGGGCAACAGATGAGACCTGTTTCTTAAAGAAAAAGAAAA 3510	

Query Match	8.8%	Score 380.2	DB 2	Length 183658
Best Local Similarity	62.2%	Pred. No. 1.3e-69		
Matches 697	Conservative	1	Mismatches 389	Indels 34
Gaps				
Qy 2391	TCATCTCAAAAATAAAACCATAGGCGGCGGCTGGCTACGCGCTGTAAATCCACAGCAC	2450		
Db 21772	TCAGTCAAAATAAAGGCTGGAGCGGCGGCTGGCTACGCGCTGTAAATCCACAGCAC	21831		
Qy 2451	TTTGGGAGCCGAGCGGCTGGATCATGAGTCAAGAGATCGAGACCATCTCGGTAAACA	2510		
Db 21932	TTTGGGAGCCGAGCGGCTGGATCATGAGTCAAGAGATCGAGACCATCTCGGTAAACA	21891		
Qy 2511	AGGTGAACCCCGCTCTACTAAAAATACAAAAATAGCGGCGCGGTGGCGGCGGCC	2570		
Db 21892	AGGTGAACCCCATCTCTATTAAAAATACAAAAATAGCGGCGCGGTGGCGGCGGCC	21951		
Qy 2571	TGTAGTCCAGCTCTCCGGAGGCTGAGCGAGGAGATGCGGTCAACCCGGGAGCGGAG	2630		
Db 21952	TGTAGTCCAGCTCTCCGGAGGCTGAGCGAGGAGATGCGGTCAACCCGGGAGCGGAG	22011		
Qy 2631	CTTGCACTGAGCAAGATTGCGCCACTCGAGTCCGCGAGTCCGCGCTGGGCGACAGACGA	2690		
Db 22012	CTTGCACTGAGCGAGATTGCGCCACTCGAGTCCGCGAGTCCGCGCTGGGCGACAGACGA	22071		
Qy 2691	GACTCTGTTCAAAAAAATAAAAAAATAAAAAACCATAAATGAGGAACCCATCTTT	2750		
Db 22072	GACTCTGTTCAAAAAAATAAAAAAATAAAAAACCATAAATGAGGAACCCATCTTT	22131		
Qy 2751	ACACTTAGGGTTTGAGTTTCTGTATCTATAAAAAAGGTTTGGATTAAAGTGATCCCTGGC	2810		
Db 22132	T-----TCAGGAGCGGGTTTTTACTCGGTTTTTCAATGTT	22166		
Qy 2811	ACTTATAAATGTTAGGCGCTTAATATTATCATAGTCGAGGATAGTTTTCATCTTTAGTC	2870		
Db 22167	ACATCCACATGCTACACAGCATCTCGACATTTTAGTTCATGCGATGCGATTTA	22226		
Qy 2871	GCCTCCTTAGTCACCTCTTCCTATACCAATCTGAGACCATTTACATTTAGAAAAACAA	2930		
Db 22227	GTGAATTGAATTAATTTCTTGAATAAAGAAAGGTATAAGATTAAATATACAAAGTTT	22286		
Qy 2931	ATAACTGGTGGGTACTTGATAGTATATAACCAAGAAAAATAATTTTAGAAGGAATTA	2990		
Db 22287	CAATGGGACTATGGGAAGCAATGAAGACCTTTTCTTTTTTTTTTTCATGACCTTG	22346		
Qy 2991	AGTTTGAACCCACATGTTAACAATACTTACCAAGTGGGATTTGCCTGTGATTAAGATG	3050		
Db 22347	TCGGATATATCCATAAAAAAATTCATTAATGTTGAGGGAATAAATAAATATTAAG	22406		
Qy 3051	CTGTAACATTTGGGCGAGTAGTTATAATTTGAAAAATGTTTATAGCAATATATAATTT	3110		
Db 22407	CTTTCAAAAAATTTAAAAACCATGTACACAGAAAAATAAAAAATGCATTAGCACCAGAT	22466		
Qy 3111	TTTATTAAATATACAGTTTTCATCACTATAGTATTTCATTAAGTCTAAGATGCCATC	3170		
Db 22467	ATTGTGTCAGTGAAT--TAAATTTCCGCTTAAGTATATATTAAGCATTTGTGAACATG	22524		
Qy 3171	AGTGGTTAGCAACACCACTGTTTTATGCATGCTTAAGAAAGATAAAGGCTGTGTGCA	3230		
Db 22525	ATCACATCGCAAGTGACAAAAATTCGAATAAATTTGATAAAGTAGTGTCTCACACCTGTA	22584		
Qy 3231	GTGGCTCACACCTGTGGGAGCCAGGACAGGACATCACTTGAGGCCAGAGTTCAAGAC	3290		
Db 22585	AT--CCCATCACTTTGGGAGCCAGGCGGCGACATCACTTGAGGTGAGGATTTAAAC	22642		
Qy 3291	CACCTCTGGTCAACATTGTAAAGCCCTGTCTCTACAAAAAATAAAGATTTAAAAATAGCT	3350		
Db 22643	CAGCCTGGTCAACATGGTGAAGTCTGTGTTCTACCAAAAAATA--CAAAAAATTAGCT	22698		
Qy 3351	GGTGGGTTGGACATGCGCTGTAGTTCAGCTACTCTGGAGGCTAAGGTGGGAGGATTCG	3410		
Db 22699	GGGTGGTGGTGCATGCGCTGTAGTCCAGCTAGTTGGGCACTGTGTGGGAGAAATTCG	22758		
Qy 3411	TAGAGCCACGGTGTGGAGCTGCAAT--TGAGCTGTGACCACACCACTGCGCTCCAGCGTG	3469		

Qy 2408 ACCATAGCGCGGGTGGCTCAGCCTGTAATCCAGCACCTTTGGGAGGCCGAGCG 2467
 |||||
 Db 107646 TAATGTGCGCGCGGGTGGCTCAGCCTGTAATCCAGCACCTTTGGGAGGCCGAGCG 107587
 |||||
 Qy 2468 GGTGGATCATGAGTCAGGAGATCGAGACCATCTGGCTTAACAAGGTGAACCCCGTCTC 2527
 |||||
 Db 107586 GGTGGATCATGAGTCAGGAGATCGAGACCATCTGGCTTAACAAGGTGAACCCCGTCTC 107527
 |||||
 Qy 2528 TACTTAAATAACAAAAATTTAGCGCGCGGGTGGCGGCGCTGTAGTCCAGCTTCTC 2587
 |||||
 Db 107526 TACTTAAATAACAAAAATTTAGCGCGCGGGTGGCGGCGCTGTAGTCCAGCTACTC 107467
 |||||
 Qy 2588 GGGAGGCTGAGGAGGAGAGATGCGGTGAACCGGAGCGGAGCTTGCAGTGAGCCAGA 2647
 |||||
 Db 107466 GGGAGGCTGAGGAGGAGAGATGCGGTGAACCGGAGCGGAGCTTGCAGTGAGCCAGA 107407
 |||||
 Qy 2648 TTGGCGCACCTGCAGTCGGCAGTCGGCCTGGCGGACAGAGCGAGACTGTGTCAAAAAA 2707
 |||||
 Db 107406 TTGGCGCACCTGCAGTCGGCAGTCGGCCTGGCGGACAGAGCGAGACTCGTCTCAAAAA 107347
 |||||
 Qy 2708 AAAAAAATAAAAAAACCAATAATGAGGAACCGCATCTTTACACTTAGGGTTTGTAGT 2767
 |||||
 Db 107346 AAAAAAATAAAAAAACCAATAATGAGGAACCGCATCTTTACACTTAGGGTTTGTAGT 107287
 |||||
 Qy 2768 TTCTGTATCTATAAAAAAGGTTTGGATTAAAGTGATCCCTGGCACCTTAAATGTTAGG 2827
 |||||
 Db 107286 ATATATATATCTTAAATGTTCTTAAATGATTGATAGACACATTTCAACATTTTCTCT 107227
 |||||
 Qy 2828 GCTTAAATATT-----ATTCTATAGATCGAGGATATCTTCAATCTTA 2867
 |||||
 Db 107226 ATCTCTCTTTCACAAGTCTCTTTTCCAAATCAGAGGAAGTTCTACTTCTCTCATAT 107167
 |||||
 Qy 2868 GTCGCCCTCTTAGTCACCTCTCTATACCAATCTGAGACCATTTTACATTTAGAAAAA 2927
 |||||
 Db 107166 GGTTCATGGAATTTGCCATCATTTCTTTTGGCATCTGTATCAGTTTGTAGGGTTGC 107107
 |||||
 Qy 2928 CAAATAACTGTTGGGTACTTGTATAGTATAATAAACAAGAAAAATATTTTAGAAGAA 2987
 |||||
 Db 107106 CATTACAAAGCACTACAAACTTGTAGCTTAAGCAACAGAAATTTAATCTCCACAGTTC 107047
 |||||
 Qy 2988 TTAAGTTTGAACACATGTTTACAAATTTACCAAGTGGGATTTGCCCTGTGATTAAAG 3047
 |||||
 Db 107046 TAGAGGCTGGAAGTCCCAAGATCAAAATGTTGGCATGACCTGGTTTCTGTAGCGCTTCC 106987
 |||||
 Qy 3048 ATGCTGTAAACATTTGGGCGAGTAGTTATATTTTGAATAATGTTTATAGCCAAATATAA 3107
 |||||
 Db 106986 TTTTGGCTTGTAGATGGGCACTTTATCTCTGTCTTACATGATCTCCCTCTCGGTG 106927
 |||||
 Qy 3108 TTTTATTTTAAATATACAGTTTTCATCAGTCTATTAGTATTTTCAATTAAGTCAAGTCC 3167
 |||||
 Db 106926 TCTGTATCTTAATCTCTACTTCTTATAAGGACACCACTATATCGAATTTAGGACCCACC 106867
 |||||
 Qy 3168 ATCAGTGTTTGAACACCACTGTTTATGCACTGCTGAAGAAGATAA----- 3217
 |||||
 Db 106866 CAATGACCTCATTTAGGCTTACTTATTTTTCAAAGTTTCTACTCTCAATTAAGTCA 106807
 |||||
 Qy 3218 --AGGCTGTGTGAGTGGCTCAACCTG-----TGGAGCCGCCAAGGAGCAG 3261
 |||||
 Db 106806 TTCTGGCTGGCGGTGGTGGCTCACACTTGTAAATCCAGAGTTTTGGAGGCCAATGTGG 106747
 |||||
 Qy 3262 AGCATCACTTGGAGCCGAGAGTTCAAGACCAACCTGTCACCATTTTAAGACCTCTCTC 3321
 |||||
 Db 106746 TGGATCACTTGGAGTCAGGAGTTTGGAGACCAAGCTGCCCCAACATATGAACCCCAATCTC 106687
 |||||
 Qy 3322 TACAAAAAATAAAGTTAAAAATTTAGCTGGGTGGCGTGGCAGCATGCTGTGTGTTCCAGC 3381
 |||||
 Db 106686 TACTTAAAAATA-----CAAAAAATTTAGCTGGCGGTGGTGTACACACCTTGAATCCAGC 106631
 |||||
 Qy 3382 TACTCTGAGGCTTAAAGTGGAGGATTTGCTAGAGCCACGCTGTTTGAAGCTGCAATGAGC 3441
 |||||
 Db 106630 TACTCAGGAGGCTGAGGAGTGAATTTGTTGAACCCAGGAGGAGGCTACAGTGAGC 106571
 |||||
 Qy 3442 TGTGACCACCACTGCGCTCCAGCGTGGCGGCAACAGAGTGGAGCCCTGTTTCTTAAAAAGAA 3501
 |||||

Db 106570 CGAGATCAGCCACCCACTCCAGCTGGTGACAGAGGAGCTCATCTCGAAAAAA 106511
 |||||
 Qy 3502 AGAAGA 3508
 |||||
 Db 106510 AAAAAA 106504
 |||||
 RESULT 15
 AC122718 170881 bp DNA linear HTG 25-MAY-2002
 LOCUS Homo sapiens chromosome 5 clone RP11-631M6, WORKING DRAFT SEQUENCE,
 DEFINITION 6 unordered pieces.
 AC122718
 AC122718.1 GI:21206281
 HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
 VERSION
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 170881)
 DOE Joint Genome Institute.
 TITLE Sequencing of Human Chromosome 5
 JOURNAL Unpublished
 2 (bases 1 to 170881)
 DOE Joint Genome Institute.
 DIRECT SUBMISSION
 SUBMITTED (25-MAY-2002) Production Sequencing Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 -----Genome Center
 Center: Joint Genome Institute
 Center Code: JGI
 Web site: http://www.jgi.doe.gov

 Project Information
 Center Project Name: 1537102
 Center clone name: RP11-631M6

 Summary Statistics
 Consensus quality: 164809 bases at least Q40
 Consensus quality: 167217 bases at least Q30
 Consensus quality: 168217 bases at least Q20
 Estimated insert size: 104870; agarose-fp estimation
 Estimated insert size: 170381; sum-of-contigs estimation
 Quality coverage: 10.06 in Q20 bases; agarose-fp estimation
 Quality coverage: 6.19 in Q20 bases; sum-of-contigs estimation.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 6 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 4603: contig of 4603 bp in length
 * 4604 4703: gap of unknown length
 * 4704 11301: contig of 6598 bp in length
 * 11302 11401: gap of unknown length
 * 11402 25882: contig of 14481 bp in length
 * 25883 25982: gap of unknown length
 * 25983 56850: contig of 30868 bp in length
 * 56851 56951: gap of unknown length
 * 56951 104415: contig of 47465 bp in length
 * 104416 104515: gap of unknown length
 * 104516 170881: contig of 66366 bp in length.
 FEATURES
 Location/Qualifiers
 1..170881
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="5"
 /clone_lib="RP11-631M6"
 /clone_lib="RP11-631M6" human BAC library 11"
 BASE COUNT 51417 a 34652 c 33285 g 50952 t 575 others

ORIGIN

Query Match 8.7%; Score 377.6; DB 2; Length 170881;
Best Local Similarity 59.8%; Pred. No. 4.6e-69; Indels 28; Gaps 5;
Matches 734; Conservative 1; Mismatches 465;

QY 2299 CCTCTCCATTTGFGATCCCTTATAGGAATAATCCTGTAATTTGTTTATGATGAGACGAGC 2358
Db 450 CCAGAACTTTTCTTATACATGCTAATACCTTTTGGATTTGGTTTAAGTAATAGGACATC 509
QY 2359 ATTGGTTTGTATATCATATCTGTGTTTCTTTCATCCTTAAAAATAAATAACCATAGCGCG 2418
Db 510 TTATTTTTTAAAAACAAAAAATATGCTAAACATCAGTAATATTTCTACTGAGAGTGGCCG 569
QY 2419 GGC CGGTGCTACGCTCTAATCCAGCACCTTGGAGCGCGAGCGGTGGATCATG 2478
Db 570 GGC CGGTGCTACGCTCTAATCCAGCACCTTGGAGCGCGAGCGGTGGATCATG 629
QY 2479 AGGTCAGGAGATCGAGACCATCTCGCTAAACAAGGTGAAACCCCGCTCTCTACTAAAAATA 2538
Db 630 AGGTCAGGAGATCGAGACCATCTCGCTAAACAAGGTGAAACCCCGCTCTCTACTAAAAATA 689
QY 2539 CAAAAAATTAGCCGGCGCGGTGGCGGCGCTGTAGTCCCGCTTCTCGGAGGCTGAG 2598
Db 690 CAAAAAATTAGCCGGCGCGGTGGCGGCGCTGTAGTCCCGCTTCTCGGAGGCTGAG 749
QY 2599 GCAGGAGATGCGGTGAACCGGGAGCGGAGCTTGCGAGTGAGCCAAAGATTGCGCCACTG 2658
Db 750 GCAGGAGATGCGGTGAACCGGGAGCGGAGCTTGCGAGTGAGCCAAAGATTGCGCCACTG 809
QY 2659 CAGTCCGCGAGTCCGCGCTGGCGAGCAGAGCGAGACTCTGTCTCAAAAAAATAAAAAA 2718
Db 810 CAGTCCGCGAGTCCGCGCTGGCGAGCAGAGCGAGACTCCGTCTCAAAAAAATAAAAAA 869
QY 2719 AAAAAA-----ACNTAATAGGAACGCACTCTTACACTTAGGGTTT 2763
Db 870 AAAAAATTTCTACTGAGAAACACTACTACTTTTACACACCTAAGAAATTTCCGCAGATA 929
QY 2764 GAGTTTCTGTATCTATAAAAAAGGTTTGGATTAAGTGATCCCTGGCACCTTATAAAATGT 2823
Db 930 AATGTTCTGCTTCTTGTCTAAACAGAGCTCTTTGATCTTTGCTGGTAAGNAITCTTCTG 989
QY 2824 TAGGCGTTTAATATTATTCATAGATCGAGGATAGTCTTCAATTTAGTCGCGCTCTTAGTCA 2883
Db 990 AACTGAATAAAGTTTTTGAACATGCTAGTGGTGAAGAGTATACATAGCCCTTTGAGT 1049
QY 2884 CTCCTCTATACCATCTGAGACCAATTTTACAATTTAGAAAAGACAATAACTGTTGGG 2943
Db 1050 CTGGCGAGGGAAGAAAGAAAGTTTCAAAAGGCTAAACAAAGCCACCTCTCATAAAC 1109
QY 2944 TTACTTTGATAGTATAAACAAGAAAAATAATTTTAGAAGAAATTAAGTTTGAACCCAC 3003
Db 1110 CTGATGAACAGTTGAGTCTGAATTCATGAGATTATGGGGTTTAGAAAGTGGAAGATGGA 1169
QY 3004 ATGTTTAACAAATTTACCAAGTGGGATTTGGCTGTGATTAAGATGCTGTAACATTTG 3063
Db 1170 TAGGCGACCACTTTGGCTTTAGTGAGATTAATGTTCTGTTGCTTGGATAAAACTCAGG 1229
QY 3064 GCGCAGTAGTATAATTTGAAAATGTTTATAGCCAATATATAATTTTATTTAAATAT 3123
Db 1230 TGCTGGTGGGGTGAATTTGATGGGAATATAGAACTTTTGTGTCGCCCAACTG 1289
QY 3124 ACAGTTTCATGAGTCTATTAGTATTTCATTA--GTCTAAGATGCCATCAGTGGTTAGCA 3181
Db 1290 CTTAATCTTAGTCTGTGGACCATATGTAATGSCCTTCTTTTCATATTCACCTTTTTT 1349
QY 3182 AACACCACTGTTTTATGCACTCTAAGAAAGAAATAAGGGCTGTGTGCACTGGCT----- 3236
Db 1350 TGCTCTGCCATCTGTGTAATAATCTTGGAGAGCGCGCGGTGGTGCACACCTGTAAAT 1409
QY 3237 --CACACCTGTGGAGCGCAAGCAGGAGCATCACTTGAGGCGCAGAGTTCAAGACCAAC 3294
Db 1410 CCCAGCACCTTTGGAGCGCGCAGGCGGATCAC--GAGGTCAGGAGATCGAGACCATC 1467

Search completed: February 28, 2003, 09:06:10
Job time : 12406 secs

QY 3295 CTGGTCAACATTTGTAAGACCCTGTCTCTACAAAAAATAAAAGATTAAAAATTAGCTGGGT 3354
Db 1468 CTGGCTAACATGGTGAAATCCCGTCTCTCTACTAAAAATACAAA--AAATAATTAACTGGGT 1525
QY 3355 GCGGTGGCACATGCCCTGTAGTTCAGCTACTCTGGAGGCTAAGGTGGGAGGATTGCTAGA 3414
Db 1526 GTGGCGGCGAGCACCTGTAGTCCCGAGCTACTCAGGAGGCTGAGGCGAGGAATGCGCTGA 1585
QY 3415 GCACCGGTGTTGGAAGCTGCAATGAGCTGTGACCAACACCACTGCGCTCCAGCGTGGGCAA 3474
Db 1586 ACCCGGGAGGCGAGGTTGCACTGAGCTGAGATCGTCCACTGCACCTCAACCTGGGAGA 1645
QY 3475 CAGAGTGAGACCCCTGTTTCTTAAAGAAA 3502
Db 1646 CAGAGTGAGACTCCCGTCTCAAAAAAATA 1673

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 28, 2003, 03:57:48 ; Search time 813 Seconds
(without alignments)
12030.048 Million cell updates/sec

Title: US-09-730-559B-7
Perfect score: 4343
Sequence: 1 atggagcccccacaaagcagg.....attcacacagctcccttct 4343

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	4341.8	100.0	4343	21	AAx82629 Human IgA nephropathy-associated protein; diagnosis; treatment; antisense;
2	990	22.8	2867	24	AAD31616 Human wild-type IR
3	877.2	20.2	3770	24	AAD31617 Rat wild-type IRp-
4	469.6	10.8	608	23	ABV56155 Human prostate exp
5	345.8	8.0	16511	22	AA536564 Human cardiovascular
6	336.4	7.7	237326	19	AAV57903 Hereditary haemoch
7	334.4	7.7	9840	22	AAK67429 Human immune/haema
8	327.6	7.5	10127	22	AAK78818 Human immune/haema
9	324.6	7.5	11125	22	AAK66446 Human immune/haema

C 10	322.8	7.4	235033	19	AAV57926 Hereditary haemoch
C 11	7831	7.4	7831	22	ABA19111 Human nervous syst
C 12	320.6	7.4	23075	22	ABA19112 Human nervous syst
C 13	317.4	7.3	465237	24	ABQ87681 Human oestrogen re
C 14	317.4	7.3	465237	24	ABQ87681 Human oestrogen re
C 15	314.8	7.2	57144	22	AAK79963 Human immune/haema
C 16	314.6	7.2	27579	22	AAK66484 Human immune/haema
C 17	314.2	7.2	154902	24	ABQ88198 Human osteoblast d
C 18	313	7.2	27960	22	AAK69779 Human immune/haema
C 19	313	7.2	27960	22	AAK69779 Human immune/haema
C 20	312	7.2	12355	22	AAK67628 Human immune/haema
C 21	311.8	7.2	698	22	AAK81472 Human immune/haema
C 22	311.8	7.2	38653	22	AAK84513 Human LEKTI DNA cl
C 23	311.8	7.2	368004	24	ABL57909 Human transporter
C 24	311.2	7.2	23855	22	AAI99411 Human excretory re
C 25	311.2	7.2	23855	22	AAI99411 Human excretory re
C 26	311.2	7.2	42738	22	AAK68992 Human kidney relat
C 27	311.2	7.2	49634	24	ABL68647 Kidney cancer rela
C 28	311	7.2	4534	20	AAK03042 Human IL-lra BAC c
C 29	311	7.2	14991	22	AAK33441 DNA encoding human
C 30	311	7.2	20210	22	AAK77417 Human immune/haema
C 31	311	7.2	20210	22	AAK77417 Human immune/haema
C 32	311	7.2	20210	24	ABQ66843 Human DNA for a no
C 33	310.4	7.1	1009	22	AAK03199 Human polynucleoti
C 34	310.2	7.1	3024	22	AAH17662 Human reproductive
C 35	310.2	7.1	172637	24	ABN83124 Human cDNA sequenc
C 36	310	7.1	333	22	AAK81556 Human voltage-acti
C 37	310	7.1	26241	22	AAK16222 Human immune/haema
C 38	309.8	7.1	32204	22	AAK05849 Human nervous syst
C 39	309.8	7.1	32204	23	ABL98413 Human reproductive
C 40	309.6	7.1	13873	22	AAK27622 Human testicular a
C 41	309.6	7.1	175737	24	ABK83571 DNA encoding novel
C 42	309.2	7.1	27007	22	AAK27764 Human cDNA differe
C 43	309	7.1	8639	20	AAK02995 DNA encoding novel
C 44	308.8	7.1	340	22	AAK82948 Human IL-lra BAC c
C 45	307.8	7.1	10339	22	AAK40445 Human immune/haema
C 45	307.8	7.1	10339	22	AAK40445 DNA encoding human

ALIGNMENTS

RESULT 1
AAx82629
ID AAX82629 standard; cDNA; 4343 BP.
XX
AC AAX82629;
XX
DT 10-NOV-2000 (first entry)
XX
DE Human IgA nephropathy-associated cDNA INM063-7.
XX
KW IgA nephropathy-associated protein; diagnosis; treatment; antisense;
KW human; INM063-7; ss.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT CDS 1..1032
FT /tag= a
FT /product= "INM063-7"
XX
PN WO9963085-A1.
XX
PD 09-DEC-1999.
XX
PF 28-MAY-1999; 99WO-JP02855.
XX
PR 02-JUN-1998; 98JP-0152603.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX
PI Ishiwata T, Sakurada M, Kawabata A, Nakagawa S, Nishi T, Kuga T;
Sawada S, Takei M, Shibata K, Furuya A;

XX WPI; 2000-097328/08.
DR P-PSDB: AAW90743.
XX
XX DNA sequences preferentially expressed in IgA nephropathy patients,
PT proteins encoded by them, and antibodies to those proteins -
XX
XX Claim 1; Page 106-113; 180pp; Japanese.
XX
XX This invention describes novel DNA sequences preferentially expressed in
CC IgA nephropathy patients, and DNA sequences stringently hybridizing to
CC them. Independent claims cover diagnostic reagents for IgA nephropathy
CC incorporating the antisense sequences; the treatment of IgA nephropathy
CC using the antisense sequences for mRNA inhibition; proteins associated
CC with IgA nephropathy, containing sequences encoded by the DNA sequences;
CC antibodies recognizing these proteins; the production of the proteins
CC by culture of host cells transformed with DNA encoding them; diagnostic
CC reagents for IgA nephropathy containing the antibodies; and compositions
CC for the treatment of IgA nephropathy which contain the antibodies. The
CC products of the invention can be used for the diagnosis and treatment of
CC IgA nephropathy. This sequence encodes the human IgA
CC nephropathy-associated protein INM063-7 which is described in the
XX method of the invention.
XX
SQ Sequence 4343 BP; 1321 A; 795 C; 865 G; 1359 T; 3 other;

Query Match 100.0%; Score 4341.8; DB 21; Length 4343;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4343; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGACGGCCCAAGACAGGATAGCCCTTCAGTACCTTATTGAAACATTAATGACAGT 60
Db 1 ATGGACGGCCCAAGACAGGATAGCCCTTCAGTACCTTATTGAAACATTAATGACAGT 60

Qy 61 TCACATAGAAGTCTTCGATGTATCAAACTTGGCACAAGTATGATGTTCTGCCCTTAC 120
Db 61 TCACATAGAAGTCTTCGATGTATCAAACTTGGCACAAGTATGATGTTCTGCCCTTAC 120

Qy 121 TCAATACGGGCTCTGTTGGAAGCTGCTGACGAAATTTGTATGGCTTTTAAATGAAGAAG 180
Db 121 TCAATACGGGCTCTGTTGGAAGCTGCTGACGAAATTTGTATGGCTTTTAAATGAAGAAG 180

Qy 181 GAAGATGTTATGAACATTTTACAGATTTTACTGGAATACCAGCAATGTTGAAGTCCCTTT 240
Db 181 GAAGATGTTATGAACATTTTACAGATTTTACTGGAATACCAGCAATGTTGAAGTCCCTTT 240

Qy 241 TTCCCTGCCCGTCTTCTTCAAGATTTTACTGGAATACCAGCAATGTTGAAGTCCCTTT 300
Db 241 TTCCCTGCCCGTCTTCTTCAAGATTTTACTGGAATACCAGCAATGTTGAAGTCCCTTT 300

Qy 301 GCTATGAGGGAGGCGAGTGAAGTCTTGGAGGTGATCCTGAGAAAGTCCATCCTGCTTGT 360
Db 301 GCTATGAGGGAGGCGAGTGAAGTCTTGGAGGTGATCCTGAGAAAGTCCATCCTGCTTGT 360

Qy 361 CGGACAGATCTTACAGTTGACCATTTTCAAAATGACATTCAGTAAATGTCGAATACAG 420
Db 361 CGGACAGATCTTACAGTTGACCATTTTCAAAATGACATTCAGTAAATGTCGAATACAG 420

Qy 421 AATGACCAAAATCCTGGAGGTGGTACCTGCAGAAAGCAGGAAAGCTCTCCCACTTAAA 480
Db 421 AATGACCAAAATCCTGGAGGTGGTACCTGCAGAAAGCAGGAAAGCTCTCCCACTTAAA 480

Qy 481 GTGACGCTTAAGAAGCTTCCCTGCGAGGGCCAGACTACCTCCGAGGATCTGTGATTCT 540
Db 481 GTGACGCTTAAGAAGCTTCCCTGCGAGGGCCAGACTACCTCCGAGGATCTGTGATTCT 540

Qy 541 GGAGAACTAGGCGGAAACTCAGGAACATTTTCTTCGAGATGTGAATACACCCATCCTG 600
Db 541 GGAGAACTAGGCGGAAACTCAGGAACATTTTCTTCGAGATGTGAATACACCCATCCTG 600

Qy 601 TGTCCTTTTCAATTTGCAACACAGTGCCTGAACCTGAAACAGTGTAAAAAATCAAGAAGTA 660
Db 601 TGTCCTTTTCAATTTGCAACACAGTGCCTGAACCTGAAACAGTGTAAAAAATCAAGAAGTA 660

Qy 661 GAATTCGGCAGAAATCGAGAGAGGCTTCAGTTTTTTTAACTGGAGTTCAAGAGTTTTAAAG 720
Db 661 GAATTCGGCAGAAATCGAGAGAGGCTTCAGTTTTTTTAACTGGAGTTCAAGAGTTTTAAAG 720

Qy 721 AATGTGGCAGTGATCCTCTCGGAAGTGAATGGCTCATCAAAATAACATTAAGATATTG 780
Db 721 AATGTGGCAGTGATCCTCTCGGAAGTGAATGGCTCATCAAAATAACATTAAGATATTG 780

Qy 781 TCAAGAGTGGTTTTTCAAGAAAAAGACCTCTCTCCACAGACAGTCTAGTCGGCACAGAT 840
Db 781 TCAAGAGTGGTTTTTCAAGAAAAAGACCTCTCTCCACAGACAGTCTAGTCGGCACAGAT 840

Qy 841 TCACACATAACGATGATGATGTTAGGATTTCTGGGGTGGGGGTTCGAGGCAATTGAA 900
Db 841 TCACACATAACGATGATGATGTTAGGATTTCTGGGGTGGGGGTTCGAGGCAATTGAA 900

Qy 901 ACAGAAGCAGTTATGCTTGGTCTGCCAGCTTCTCTTACTTTACCAGAGTGGTGGATGT 960
Db 901 ACAGAAGCAGTTATGCTTGGTCTGCCAGTCTCTCTTACTTTACCAGAGTGGTGGATGT 960

Qy 961 GAGTTAACTGGGTCATCAAAACCCCTTTTGTATCATCCATAGATGTTCTTCTGGTATTACA 1020
Db 961 GAGTTAACTGGGTCATCAAAACCCCTTTTGTATCATCCATAGATGTTCTTCTGGTATTACA 1020

Qy 1021 AAGGTAAAGTTAAAGTTGGTAGCTCTATGACTTACTGAACATTTATTTTATAAAAAATG 1080
Db 1021 AAGGTAAAGTTAAAGTTGGTAGCTCTATGACTTACTGAACATTTATTTTATAAAAAATG 1080

Qy 1081 AAGAGCTCTATGAGACAGGATTTGGGTTTCTTACTTCTCAGGTCCTTGACGTT 1140
Db 1081 AAGAGCTCTATGAGACAGGATTTGGGTTTCTTACTTCTCAGGTCCTTGACGTT 1140

Qy 1141 AGCCACATCATCATAGTTATCATAGTAATAACAACAACAGAGCATTTAGTTGTACTAA 1200
Db 1141 AGCCACATCATCATAGTTATCATAGTAATAACAACAACAGAGCATTTAGTTGTACTAA 1200

Qy 1201 TAAATACAAAGAAATTTGTTGTTTCCACTTATGTTAGTCAATTTAGTCTTATACAACAG 1260
Db 1201 TAAATACAAAGAAATTTGTTGTTTCCACTTATGTTAGTCAATTTAGTCTTATACAACAG 1260

Qy 1261 CTGTGAGATGGATACATTTACTTACTTCTCATTTCTGTAAGTCTGAGAAACTAAGGTACAGTAG 1320
Db 1261 CTGTGAGATGGATACATTTACTTACTTCTCATTTCTGTAAGTCTGAGAAACTAAGGTACAGTAG 1320

Qy 1321 GGTTTAGTGAATACCAAGGGTCGAAAGGCTGAGTATAAGGGTAGACAAAGATTCCA 1380
Db 1321 GGTTTAGTGAATACCAAGGGTCGAAAGGCTGAGTATAAGGGTAGACAAAGATTCCA 1380

Qy 1381 GGAGTCAGATTTCTGAGTCATGCTTAACATTTATGCCCTTATTTAGTGCCTTGTGGCTTA 1440
Db 1381 GGAGTCAGATTTCTGAGTCATGCTTAACATTTATGCCCTTATTTAGTGCCTTGTGGCTTA 1440

Qy 1441 ATAAACACTTGTGGACTACATATTTTCTCTCTTTTAACTTCAATTAATAAAAAAAT 1500
Db 1441 ATAAACACTTGTGGACTACATATTTTCTCTCTTTTAACTTCAATTAATAAAAAAAT 1500

Qy 1501 GTTTAGCAAAAGTTGAWTGTGCTCTTTAAATTAATTTATTTGCCCGTTAGAACTGTG 1560
Db 1501 GTTTAGCAAAAGTTGAWTGTGCTCTTTAAATTAATTTATTTGCCCGTTAGAACTGTG 1560

Qy 1561 CTCTACTAAGTAATGCTTTCAAAAAACATGGACTGTAGAATGTGATATATCATTTTCTG 1620
Db 1561 CTCTACTAAGTAATGCTTTCAAAAAACATGGACTGTAGAATGTGATATATCATTTTCTG 1620

Qy 1621 TTGCCCGTTTAAACATTTCTCTGATTTATGTAAAAATCTTCTCTCTGAAATTTTAAAA 1680
Db 1621 TTGCCCGTTTAAACATTTCTCTGATTTATGTAAAAATCTTCTCTCTGAAATTTTAAAA 1680

Qy 1681 TACTGGCTTCAGAACTTCAATACATACACTGAGCTTGTAAAGCATATTAAATACACAGGCT 1740
Db 1681 TACTGGCTTCAGAACTTCAATACATACACTGAGCTTGTAAAGCATATTAAATACACAGGCT 1740

QY 1741 CAGGATTTCTAGTGAACAATAATTTGTAACTCTTCTCTCTAAATGCTGGCCTTTGCT 1800
DB 1741 CAGGATTTCTAGTGAACAATAATTTGTAACTCTTCTCTCTAAATGCTGGCCTTTGCT 1800
QY 1801 AACTTATTTAAATGATTAAATCCTATTCTTTGTTAAATGAATGTACCTGGAAATGTTCCA 1860
DB 1801 AACTTATTTAAATGATTAAATCCTATTCTTTGTTAAATGAATGTACCTGGAAATGTTCCA 1860
QY 1861 CATATAATTCOAATTTGAGTCCCAATCTCAGCAATTTTGGTTAGATTATTGGTAGCAAGG 1920
DB 1861 CATATAATTCOAATTTGAGTCCCAATCTCAGCAATTTTGGTTAGATTATTGGTAGCAAGG 1920
QY 1921 CTTTCTGGATACTCCAGTGTAAAGAAATGATAATGCTCCCTCTCAGCAATTTGGTATTGA 1980
DB 1921 CTTTCTGGATACTCCAGTGTAAAGAAATGATAATGCTCCCTCTCAGCAATTTGGTATTGA 1980
QY 1981 TCCTTTCTCCCTAAATAGAAAAGAAATTTGGCATCTTAGAGAAATATTGATTCAACGTAT 2040
DB 1981 TCCTTTCTCCCTAAATAGAAAAGAAATTTGGCATCTTAGAGAAATATTGATTCAACGTAT 2040
QY 2041 GATACCAAAGATCAAGTAGTAATTTGGGAATTCGAGGATTATTCCTAGAGGAAAAGGAG 2100
DB 2041 GATACCAAAGATCAAGTAGTAATTTGGGAATTCGAGGATTATTCCTAGAGGAAAAGGAG 2100
QY 2101 TATCCCATTAATGTTTTACAGAAATCAATCTTTACTTTTAGACATCCTGAAAACCTAACGC 2160
DB 2101 TATCCCATTAATGTTTTACAGAAATCAATCTTTACTTTTAGACATCCTGAAAACCTAACGC 2160
QY 2161 TGCCTTTTAGCCTTCTAGCTGTTTTCTGACAAATATTACTGTGTGTTTTTGGACAT 2220
DB 2161 TGCCTTTTAGCCTTCTAGCTGTTTTCTGACAAATATTACTGTGTGTTTTTGGACAT 2220
QY 2221 TTTAGTTTAATGTTAAAAATTAATCTATTATATATGTTTACATTTATTGCAATATTGA 2280
DB 2221 TTTAGTTTAATGTTAAAAATTAATCTATTATATATGTTTACATTTATTGCAATATTGA 2280
QY 2281 TTACTCTCTTTTTCAGATCCTGTTCATTTGTGATCTTATAGGAATAATCCGTGATTGT 2340
DB 2281 TTACTCTCTTTTTCAGATCCTGTTCATTTGTGATCTTATAGGAATAATCCGTGATTGT 2340
QY 2341 TTTTTTGATGAGACAGCAATTTGGTTTGTAAATATCTAATCTGTGTTCTTTTCATCTCAA 2400
DB 2341 TTTTTTGATGAGACAGCAATTTGGTTTGTAAATATCTAATCTGTGTTCTTTTCATCTCAA 2400
QY 2401 AAATAAAACCATAGCCGGCGGGTGGCTCACGCCGTGTAATCCCAAGCACTTTGGAGGC 2460
DB 2401 AAATAAAACCATAGCCGGCGGGTGGCTCACGCCGTGTAATCCCAAGCACTTTGGAGGC 2460
QY 2461 CGAGCGGGTGGATCATGAGTCAGGAGATCGAGACCATCCTGGCTAACAAAGGTGAACC 2520
DB 2461 CGAGCGGGTGGATCATGAGTCAGGAGATCGAGACCATCCTGGCTAACAAAGGTGAACC 2520
QY 2521 CCGTCTCTACTAAAAATACAAAAATTAGCCGGCGGGTGGCGGGCGCCTGTAGTCCCA 2580
DB 2521 CCGTCTCTACTAAAAATACAAAAATTAGCCGGCGGGTGGCGGGCGCCTGTAGTCCCA 2580
QY 2581 GCTTCTCGGAGGCTGAGGCAGGAGATGCGGTGAACCCGGGAGGCGAGCTTGCAGTGA 2640
DB 2581 GCTTCTCGGAGGCTGAGGCAGGAGATGCGGTGAACCCGGGAGGCGAGCTTGCAGTGA 2640
QY 2641 GCCAAGATTGGCCACTGCACTCCGACAGTCAGGAGATCGAGACCATCCTGGCTAACAAAGGTGAACC 2700
DB 2641 GCCAAGATTGGCCACTGCACTCCGACAGTCAGGAGATCGAGACCATCCTGGCTAACAAAGGTGAACC 2700
QY 2701 CAAAAAATAAAAAAAAAAAAAACCAATAATGAGGAAACGCATCTTTTACACTTAGGG 2760
DB 2701 CAAAAAATAAAAAAAAAAAAAACCAATAATGAGGAAACGCATCTTTTACACTTAGGG 2760
QY 2761 TTTGAGTTTCTGATCTATAAAAAAGGGTTTGGATTAACTGATCCCTGGCACTTATAAAA 2820
DB 2761 TTTGAGTTTCTGATCTATAAAAAAGGGTTTGGATTAACTGATCCCTGGCACTTATAAAA 2820
QY 2821 TGTAGGCCCTAATATTATTTCATAGATCAGGAGTAGTTTCATCTTAGTCGCCCTCCTTAG 2880

DB 2821 TGTAGGCCCTAATATTATTTCATAGATCAGGAGTAGTTTCATCTTAGTCGCCCTCCTTAG 2880
QY 2881 TCACCTCTTCTCTATACCAATCTGAGACCAATTTACAAATTTAGAAAACACAAAATACGTGT 2940
DB 2881 TCACCTCTTCTCTATACCAATCTGAGACCAATTTACAAATTTAGAAAACACAAAATACGTGT 2940
QY 2941 GGGTTACTTGTATAGTATAATAACCAAGAAAAATAATTTTAGAAGGAATTTAACTTTGAAC 3000
DB 2941 GGGTTACTTGTATAGTATAATAACCAAGAAAAATAATTTTAGAAGGAATTTAACTTTGAAC 3000
QY 3001 CACATGTTTAAACAAATCTACCAAGTGGGATTTGCCGTGTGATTAAAGATGCTGTAACAT 3060
DB 3001 CACATGTTTAAACAAATCTACCAAGTGGGATTTGCCGTGTGATTAAAGATGCTGTAACAT 3060
QY 3061 TTGGGCCACTAGTATAAATTTGAAAAATCTTTATAGCCCAATATATAATTTTATTATTA 3120
DB 3061 TTGGGCCACTAGTATAAATTTGAAAAATCTTTATAGCCCAATATATAATTTTATTATTA 3120
QY 3121 TATACAGTTTTCATCAGTCTATTAGTATTTCATTAAGTCTTAAGATGCCATCAGTGGTTAGC 3180
DB 3121 TATACAGTTTTCATCAGTCTATTAGTATTTCATTAAGTCTTAAGATGCCATCAGTGGTTAGC 3180
QY 3181 AAACACCACCTGTTTATGCACTGCTAAGAAAATAAAGGGCTGTGTGAGTGGCTCACA 3240
DB 3181 AAACACCACCTGTTTATGCACTGCTAAGAAAATAAAGGGCTGTGTGAGTGGCTCACA 3240
QY 3241 CCTGTGGGACGCCAAGGACAGGAGCATCTTCAGGCCAGAAGTTCAAGACCACCTGGTC 3300
DB 3241 CCTGTGGGACGCCAAGGACAGGAGCATCTTCAGGCCAGAAGTTCAAGACCACCTGGTC 3300
QY 3301 AACATTGTAAAGACCCCTGCTCTACAAAAAAGTTAAAAATTAGTGGTGGCGGTG 3360
DB 3301 AACATTGTAAAGACCCCTGCTCTACAAAAAAGTTAAAAATTAGTGGTGGCGGTG 3360
QY 3361 GCACATGCCGTGATGTTCCAGCTACTCTGGAGGCTAAGGTGGGAGGATTGCTAGAGCCACG 3420
DB 3361 GCACATGCCGTGATGTTCCAGCTACTCTGGAGGCTAAGGTGGGAGGATTGCTAGAGCCACG 3420
QY 3421 GTGTTGGAAGCTGCAATGAGCTGTACCACCACTCCAGCTCCAGCTGGGCAACAGAGT 3480
DB 3421 GTGTTGGAAGCTGCAATGAGCTGTACCACCACTCCAGCTCCAGCTGGGCAACAGAGT 3480
QY 3481 GAGACCCCTGTTTCTAAAAGAAAAAGAGGGCTGCCACCTAAAACAGACACACTAT 3540
DB 3481 GAGACCCCTGTTTCTAAAAGAAAAAGAGGGCTGCCACCTAAAACAGACACACTAT 3540
QY 3541 TGAGTTGAGGTACCCCTGATTTCAAACACATGAAAATGTTTAAATATAGCCACCTTGAAGCT 3600
DB 3541 TGAGTTGAGGTACCCCTGATTTCAAACACATGAAAATGTTTAAATATAGCCACCTTGAAGCT 3600
QY 3601 TTTGAGSCCCTTCTACCTGAAATTAACAGTGACATTTGGACCAGTCTTCTCTTTACTTCT 3660
DB 3601 TTTGAGSCCCTTCTACCTGAAATTAACAGTGACATTTGGACCAGTCTTCTCTTTACTTCT 3660
QY 3661 TATCTTAAATACCCCAACCAAGATGAGTTGATTATAGGACAAATGAAGGATCTCA 3720
DB 3661 TATCTTAAATACCCCAACCAAGATGAGTTGATTATAGGACAAATGAAGGATCTCA 3720
QY 3721 TTTCCCTTCCACTCAGTAGTATTGGTTTAAAAATTTTATTTATAGTTTTCAGACAAATCGTT 3780
DB 3721 TTTCCCTTCCACTCAGTAGTATTGGTTTAAAAATTTTATTTATAGTTTTCAGACAAATCGTT 3780
QY 3781 GCTAATCTTATCTTTGCAATTTTGTATGTTTCTGTTCTTATATAGCCACTCAG 3840
DB 3781 GCTAATCTTATCTTTGCAATTTTGTATGTTTCTGTTCTTATATAGCCACTCAG 3840
QY 3841 GCAAGTAGGAGTGGCTGAAAGTTTGTTCAGTTTTTTTGGAAAGTGGAGTTTACAAATATC 3900
DB 3841 GCAAGTAGGAGTGGCTGAAAGTTTGTTCAGTTTTTTTGGAAAGTGGAGTTTACAAATATC 3900
QY 3901 TATAGTTGATCGAACTACAATAGCAAAACATGTGTCGGGAATATGGTGTATCCTCAGCTT 3960
DB 3901 TATAGTTGATCGAACTACAATAGCAAAACATGTGTCGGGAATATGGTGTATCCTCAGCTT 3960

Qy	958	TTGTGAGTTAACTGGGTCATCAAAACCCCTTTTGTTCATCATCCATAGATGTTGCTTCTGGGTAATT	101
Db	972	TTGTGAGCTAACTGGGTCATCCCAACGCTTTTGTTCATCCATAGATATTGTCTAGGCATT	1031
Qy	1018	ACAAAG 1023	
Db	1032	ACAAAG 1037	
RESULT 4			
ID	ABV56155	standard; cDNA; 608 BP.	
XX	AC		
XX	AC		
XX	ABV56155;		
DT	17-SEP-2002	(first entry)	
XX	XX		
DE	XX	Human prostate expression marker cDNA 56146.	
XX	XX		
KW	Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker		
KW	pharmacogenomic marker; gene; ss.		
OS	Homo sapiens.		
XX	XX		
PN	WO200160860-A2.		
XX	XX		
PD	23-AUG-2001.		
XX	XX		
PF	20-FEB-2001; 2001WO-0505171.		
XX	XX		
PR	17-FEB-2000; 2000US-183319P.		
PR	16-MAR-2000; 2000US-189862P.		
PR	25-MAY-2000; 2000US-207454P.		
PR	09-JUN-2000; 2000US-211314P.		
PR	18-JUL-2000; 2000US-219007P.		
PR	13-DEC-2000; 2000US-255281P.		
XX	XX		
PA	(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.		
XX	XX		
PI	Schlegel R, Endege WO, Monahan JE;		
PI	WPI; 2001-662795/76.		
DR	XX		
PT	Novel isolated nucleic acid molecule associated with cancerous state of		
PT	prostate cells and correlating with presence of prostate cancer, useful		
PT	for detecting presence of prostate cancer, stage of prostate cancer -		
XX	XX		
PS	Claim 1; Page 10839; 11750pp; English.		
XX	XX		
CC	The invention relates to an isolated nucleic acid molecule (I) compris-		
CC	a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the		
CC	specification or its complement. (I) is useful for:		
CC	(a) assessing whether a patient is afflicted with prostate cancer;		
CC	(b) monitoring the progression of prostate cancer in a patient;		
CC	(c) assessing the efficacy of a test compound to inhibit prostate		
CC	cancer in a patient;		
CC	(d) assessing the efficacy of a therapy for inhibiting prostate cancer		
CC	in a patient;		
CC	(e) selecting a composition for inhibiting prostate cancer in a patient		
CC	(f) assessing the prostate cell carcinogenic potential of a compound;		
CC	(g) determining whether prostate cancer has metastasized in a patient;		
CC	(h) assessing the aggressiveness or indolence of prostate cancer in a		
CC	patient;		
CC	(I) is also useful as a pharmacodynamic or pharmacogenomic marker.		
XX	XX		
SQ	Sequence 608 BP; 191 A; 118 C; 113 G; 185 T; 1 other;		
Query Match			
		10.8%;	Score 469.6; DB 23; Length 608;
Best Local Similarity		97.8%;	Pred. No. 1.9e-79;
Matches 497; Conservative		0;	Mismatches Indels 2; Gaps
Qy	3838	CAGGCAAGTAGGAGTGGCTGGAAAGTTTGTTCAGTTTTTTTGGAGGTGGAGTTTCACAATT	3897

PR 07-JUN-2000; 2000US-0209457.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0233080.
PR 08-SEP-2000; 2000US-0233081.
PR 12-SEP-2000; 2000US-02331968.
PR 14-SEP-2000; 2000US-0233297.
PR 14-SEP-2000; 2000US-0233298.
PR 14-SEP-2000; 2000US-0233299.
PR 14-SEP-2000; 2000US-0233400.
PR 14-SEP-2000; 2000US-0233401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 12-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0251989.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;
WPI; 2001-451930/48.
New cardiovascular system related polynucleotides and polypeptides, which encode the cardiovascular system antigen polypeptides of the invention.
Cardiovascular system antigens and their associated polynucleotides are useful for diagnosing, treating and/or preventing disorders of the cardiovascular system -
Claim 1; SEQ ID No 2064; 674pp; English.
Sequences AAS35741-AAS36942 represent genomic DNA molecules, which encode the cardiovascular system antigen polypeptides of the invention.
Cardiovascular system antigens and their associated polynucleotides are useful in the diagnosis, treatment and prevention of various types of disorders in e.g. humans, mice, rabbits, goats, horses, cats, dogs,

CC chickens or sheep. A pathological condition can be determined by
CC detecting the presence or absence of a mutation in a cardiovascular
CC system antigen polynucleotide. The treatable disorders include autoimmune
CC diseases such as rheumatoid arthritis, hyperproliferative disorders such
CC as neoplasms of the breast or liver, cardiovascular disorders such as
CC cardiac arrest, cerebrovascular disorders such as cerebral ischemia,
CC nervous system disorders such as Alzheimer's disease, infections caused
CC by bacteria, viruses and fungi, ocular disorders such as corneal
CC infection, endocrine disorders such as premature labour and infertility,
CC gastrointestinal disorders such as Crohn's disease, renal disorders such
CC as glomerulonephritis and respiratory disorders such as asthma and
CC pleurisy. The polypeptides can also be used to aid wound healing, to
CC prevent skin aging due to sunburn, to maintain organs before
CC transplantation, to regenerate tissues and in chemotaxis.
CC Note: the sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

Query Match 8.0%; Score 345.8; DB 22; Length 16511;
Best Local Similarity 61.3%; Pred. No. 8.1e-56;
Matches 682; Conservative 1; Mismatches 408; Indels 22; Gaps 7;

QY 2404 TAAACCATAGCGCGCGGTGCTCACGCCCTGTATCCAGCACCTTTGGGAGGCCGA 2463
DB 645 TTAATAAAGGCGCGGTGCTCACACCTGTATCCAGCACCTTTGGGAGGCCGA 704
QY 2464 GCGGGTGTGATCATGAGTCCAGGATCGAGACCATCTCGGTCTAACAGGTGAACCCCG 2523
DB 705 GCGGCGCATACGAGGTGCAAGATCAAAACCATCTCTGGCTACACGGTGAACCCCG 764
QY 2524 TCTCTACTAAAAATACAAA--AAATAGCGCGCGGTGCGCGCGCTGTAGTCCCGAG 2581
DB 765 TCTCTACTAAAAATGCAAAACAAATTAGCCAGGCATGTTGGTGGCAGCTGTAGTCCCGAG 824
QY 2582 CTTCTCGGAGGTGAGGAGGAGGAGTGGCGTGAACCCCGGAGGCGGAGCTTGCAGTGAG 2641
DB 825 CTACTCGGAAGTGAAGGAGGAGGAGTGGCGTGAACCCCGGAGGAGCTTGCAGTGAG 884
QY 2642 CCAAGATTCGGCACACGCTCGCGAGTCCGCGCTGGGCGAGAGCGAGCTCTGTTC 2701
DB 885 CCAAGATGGCACCACT-----GCACCTCGAGCTGGGCGAGAGCACTCCGCTC 937
QY 2702 AAAAAAATAAAAAAATAAAAAATTAATAGAGAAACGCATCTTTACACTTAGGT 2761
DB 938 AAAAAAATAAAAAAATAAAAAATTAATAAATAAAAAATTAATAAATAAAAA 997
QY 2762 TTGAGTTTCTGTATCTATAAAGAGGT--TTGGATTAAGTATCCCTGGCATTATAA 2819
DB 998 GAGAAAAAATAAAAAATGAACAAATTTGTAGTAAAAAATAAACAAAACTACCTATGGC 1057
QY 2820 ATGTTAGGGCTTAATATTATTCATAGATCGAGGATAGTTTCTTCTAGTCGCTCCTTA 2879
DB 1058 CAATAACATGGAATAAATTTCTCAGAAATATCAAAAGTACAAATGAACATTTA 1117
QY 2880 GTCACCTCTCTATACAACTCTGAGACCAATTTTACAAATTTAGAAAGACAAATACTGT 2939
DB 1118 GATAGTTTTTCATCTACCAAAATGGAAGCATATCAAAAAAATAGTAACCTGGTCACAAATGT 1177
QY 2940 TGGGTTACTTTAGTATATATACCAAGAAAAATATTTAGAGCAATTAAGTTTGA 2999
DB 1178 AGAGAGACTCTCAGTGAAGTCTCAATTTGGTTTACCGTTTCTAGACAAATTTGAGCAGTTCT 1237
QY 3000 CCACATGTTAAACAAATCTACCAAGTGGGATTTGCTGTGATTAAGATGCTGTAACA 3059
DB 1238 ATTAGTATTAATAAGTCCATGATCTATTCTGTGACCTATCTACTATTTTCCCT-GTCA 1296
QY 3060 TTTGGGCGAGTATTAATAATTTGAAAAATGTTTATAGCAATATATAATTTTATTTAA 3119
DB 1297 TTTCTGACTTTAATCAAACTCCCAAAATGTTTGAACCTTTCTATGCGGAGCTGTGAAAAAG 1356
QY 3120 ATATACAGTTTCATAGTCTATTATTAATTAAGTCTATAGATCCCATCAGTGGTTAG 3179

DB 1357 GTATGAATATTGACTAACAAATTTGTAAAGTTGAGATGCTATGAGAATA-----TGGGAA 1412
QY 3180 CAAACACCACTGTTTATGACCTGCTAAAGAAAGAAATAAAGGGCTGTGTGAGTGGCTCAC 3239
DB 1413 AATTGACATTTCTTAAGTATTATTGACTGACACAGTGGTTTACGCTGTAAATCCAG 1472
QY 3240 ACCTGTGGGACGCCAAGGACGAGCATCTCTGAGGCCAGAGATTCAAGACCAACCTGGT 3299
DB 1473 CACTTTGGGAGGCCAAGATGGGAGGATCACTTGAGGCCAGGAGTTTCAAGAACTACCTGGG 1532
QY 3300 CAACATTTGTAAGCCCTGCTCTACAAAAAATAAAGTTTAAATAATAGTGGTGGT 3359
DB 1533 CAACATAGTGAGA--CCTGTCTCAACAAAAAGTAAA-----AAAAATTAAGTGGT 1586
QY 3360 GGACATGCTGTAGTTCCAGCTACTCTGGAGGCTAAGGTGGGAGGATTTGCTAGAGCCAC 3419
DB 1587 GGTGATGCTGTGTGCTCCAGCTACTTTGGAAGGCTGAGGATGAGGATGCTTTGAGACTA 1646
QY 3420 GGTGTTGGAAGCTGCAATGAGCTGTGACACACACACTGCGCTCCAGGTGGGCAACAGAG 3479
DB 1647 GGAGATCAAGGCTGCAGTGCAGCTGTGATCATGCACTGCACCTCCAGCTGGGCAACAGAA 1706
QY 3480 TGAGACCTCTGTTCTTAAAGAAAGAAAGAAAAA 3512
DB 1707 TGAGACCTTTCTCTAAATAAATAAATA 1739
RESULT 6
AAV57903/c
ID AAV57903 standard; DNA: 237326 BP.
XX AC
XX AC AAV57903;
XX 21-DEC-1998 (first entry)
XX
DE Hereditary haemochromatosis subregion from an HH affected individual.
XX
KW Bovine butyrophilin; BT; human hereditary haemochromatosis; HFE;
KW diagnosis; iron metabolism; NPT3; NPT4; Rofet; BTF1; BTF3;
KW BTF4; BTF5; milk protein; lupus; Sjogren's syndrome; hypophosphatemia;
KW type 1 sodium transport gene; ss.
XX
OS Homo sapiens.
XX
PN WO9814466-A1.
XX
PD 09-APR-1998.
XX
PP 30-SEP-1997; 97WO-US17658.
XX
PR 07-MAY-1997; 97US-0852495.
XX
PR 01-OCT-1996; 96US-0724394.
XX
PA (PROG-) PROGENITOR INC.
XX
PI Feder JN, Kronmal GS, Lauer PM, Ruddy DA, Thomas WJ;
PI Tsuchihashi Z, Wolff RK;
XX
DR WPI; 1998-240014/21.
XX
PT Hereditary haemochromatosis gene products - used to develop products
PT for the diagnosis and treatment of hereditary disorders in iron
PT metabolism
XX
PS Claim 1; Fig 9; 209pp; English.
XX
CC The present invention describes hereditary haemochromatosis gene
CC products from the human haemochromatosis gene. The present sequence
CC represents a hereditary haemochromatosis subregion from an hereditary
CC haemochromatosis (HH) affected individual. Also described is a
CC method to determine the presence or absence of the common hereditary
CC haemochromatosis (HFE) gene mutation in an individual comprising:
CC (a) providing DNA or RNA from the individual; and (b) assessing the

CC DNA or RNA for the presence or absence of a haplotype or genotype where
CC the presence or absence of the haplotype genotype indicates the likely
CC presence of the HFE gene mutation in the genome of the individual. The
CC HFE gene sequences from the present invention can be used to develop
CC products for use in the diagnosis and treatment of hFE. The present
CC invention also describes Bf genes, which are homologues of the milk
CC protein butyrophilin (BT), and can be used in the production of agonists
CC and antagonists of Bf function. Also described are: (1) a Roret gene
CC which can be used to develop products for the study, diagnosis and
CC treatment of lupus and Sjogren's syndrome; and (2) NPT3 and NPT4 genes
CC which are homologues of a type 1 sodium transport gene, and can
CC similarly be used for hypophosphatemia.
XX
SQ

Sequence 237326 BP; 69596 A; 48904 C; 48217 G; 70609 T; 0 other;
Query Match 7.7%; Score 336.4; DB 19; Length 237326;
Best Local Similarity 59.8%; Pred. No. 8e-54;
Matches 680; Conservative 1; Mismatches 432; Indels 25; Gaps 6;

QY 2386 TTCTTTTCATCTAAAAATAAACCATAGCCGGCGCGGTGGCTCAGCGCTCTAATCC 2445
Db 137745 TTATATTTTCTTGGATAAAGTCAGTAGCGGGCGCGGTGCTCATGCTGTGATCCT 137686
QY 2446 AGCACTTTGGAGCCGAGCGGGGTGGATCATGAGTTCAGGATCGAGACCATCTCTGGC 2505
Db 137685 AGCACTTTGGAGGCTGAGGTGGCGAGATCATGAGTTCAGGATCGAGACCATCTCCGGC 137626
QY 2506 TAACAAGTGAACCCCGTCTCTACTATAAATAACAAAAATTAGCCGGCGGGTGGCGG 2565
Db 137625 TAACACGGTGAACCCCGTCTCTACTATAAATAACAAAAATTAGCTGGCGGTGGCGG 137566
QY 2566 GCGCTCTAGTCCAGCTTCTCGGAGGCTGAGCGAGGAATGGCTGAACCCGGGAGG 2625
Db 137565 GCGCTCTGTGTCAGCTACTCGGAGTTGAGCGAGGAATGGCGTGAATCCGGGAGG 137506
QY 2626 CGAGCTTTCAGTTCAGCAAGATTCGCCACTGCATCCGCGCTCGGCGGACAG 2685
Db 137505 CGGAGTTGTCAGTTCAGCGAGATGTGCCACT-----GCATTCAGCTGGCGGACAG 137453
QY 2686 AGCGAGACTCTGTCTCAAAAAAATAAAAAAATAAACCAATTAATGAGGAACGCA 2745
Db 137452 AGCAAGCTCCGCTCAGAAAAAGAAAAAAGTCAAGTAAATTTAAGAGAAAAA 137393
QY 2746 TCTTTACACTTAGGGTTGAGTTCTGTATCTATAAAAAAGGTTTGGATTAAAGTATCC 2805
Db 137392 TGCATTTGCTTTGGGACTTTTAATATTTAGTCTACAATCTAGCCACCATAGAACTCG 137333
QY 2806 CTGGCACTTATAAATGTTAGGCTTAATATTTATTCATAGATCGGAGATGTTCAATCT 2865
Db 137332 TGATTAATACGGGTCTCTGTAAATGGAAACATGCAATTTTGGGGGAAAAAGAGGAGT 137273
QY 2866 TAGTCGCTCTCTTAGTCACCTCTCTTATACCAATCTGAGACCATTTTACAATTTAGAAA 2925
Db 137272 GTTTTAGTGTATTTGTTTTTACACTGTTTATAATAAATTTAAGCAATCTTGAGGGG 137213
QY 2926 GACAATACTGTTGGGTACTGTAGTATATAAACAAGAAAAAATAATTTAGAAGG 2985
Db 137212 AACATTTTATTTCTACTTGTACTGCAATAAGTTATGAGATAAGTTTACAACCTATATCA 137153
QY 2986 AATTAAGTTTGAACACATG--TTAACAATAATCTACCAAGTGG-GAATTCGCTGTGAT 3042
Db 137152 CATACAGTTTGTAGCTTTTAAATTTATGAAATTTTAAACAGAATAATATGCTTAATATGAT 137093
QY 3043 TAAAGATGCTGTAACACATTTGGGCCAGTAGTTATATTTGAAAAATGTTTATAGCCAATA 3102
Db 137092 GAAATGTCATAAATTCATATAGATATATTTTATAAACCAATTCAGAAAGGAGCAATA 137033
QY 3103 TATAATTTTTTATTTAAATATACAGTTTTCATCAGTCTA-----TTAGTATTTTCATTAA 3155
Db 137032 CCCAATTTCAAAATCATATTAATTTGTAATTTAATTAGGCGACCAAAATATTTCTGGAAAT 136973
QY 3156 GTCTAGATGCCATCAGTGGTTAGCAAAACACACACTGTTTTTATGCACTGCTAAGAAGAAAT 3215

Db 136972 TCTTTCTAATAAAACAAATGAGTGTAAATACAGTCGTAAGTCTGACAAATCTGAAGAATTTATG 136913
QY 3216 AAAGGCTGTGTCAGTGGCTCACACCTGTGGAGCGCCAAAGCAGGAGCATCACTTGAGG 3275
Db 136912 CAGCATATAAAGTGTATATCCAGCACTTTGGGAGGCCAAAGTGGCGAGATCA--TGAGG 136855
QY 3276 CCAGAAAGTTCAAGACCAACCTGTCACAACTTCTAAGACCCCTGTCTCTACAAAAAATAA 3335
Db 136854 TCAGGAGTTGGAGACAGCCCTGACCAACATGTTGTAACCCCGCTCTTACTAAAAATA--- 136798
QY 3336 AGTTAAAAATTAGCTGGGTGGCGGTGGCACATGCCCTGTAGTTCAGTCTCTGGAGGCTA 3395
Db 136797 ---CAAAATTAGCCGGGCTTGGTGTGCACACCTGTATCCAGCTGCTCAGGAGGCTA 136741
QY 3396 AGTGGGAGGATGTAGAGCCAGCCAGCGGTGTGAAAGCTGCAATGAGCTGTGACCAACAC 3455
Db 136740 AGGCAGGAGAACTGCTTGAATCCAGGAGGTGGAAGTTGCAGTAAGCAAGATCACGCCAC 136681
QY 3456 TCGCTCCAGCGTGGGCAACAGAGTGAGACCCCTGTTTCTAAAAAGAAAGAAAAA 3513
Db 136680 TGCATCCAGCCCTGGCGACAGAGTGTGACTCTGACTCTGACTCAAAAAAATAAATA 136623
RESULT 7
AAK67429
ID AAK67429 standard; DNA; 9840 BP.
XX
AC AAK67429;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:22241.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.

PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232297.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.

PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0256179.
 PR 08-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 05-JAN-2001; 2000US-0254097.
 PR 05-JAN-2001; 2000US-0259678.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Rosen CA, Barash SC, Ruben SM;
 PI WPI; 2001-483426/52.
 XX
 DR Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX
 PS Disclosure; SEQ ID NO 33630; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 10127 BP; 2766 A; 1847 C; 2133 G; 3381 T; 0 other;
 Query Match 7.5%; Score 327.6; DB 22; Length 10127;
 Best Local Similarity 60.3%; Pred. No. 2e-52;
 Matches 689; Conservative 1; Mismatches 400; Indels 52; Gaps 7;
 QY 2405 AAACCATAGCCGGCGGTGGCTCACGCCCTGTATCCAGCACTTTGGGAGCCGAG 2464
 Db AGAATAGTGGCGGGTGGCGTGGCTTACACCTGTATCCAGCACTTTGGGAGCCGAG 5149
 QY 2465 GCGGTGGATCATGAGTTCAGGAGATCGAGACCATCTGGCTAACAGGTGAACCCCGT 2524
 Db T---CTATCACTTGGTTCAGGAGATCGAGACCATCTGGCTAACAGGTGAACCCCGT 5092
 QY 2525 CTCTACTAAATAACAAAATTAGCCGGCGGTGGCGGGCCCTGTAGTCCAGCTT 2584
 Db CTCTACTAAATAACAAAATTAGCCGGCGGTGGCGGGCCCTGTAGTCCAGCTT 5032
 QY 2585 CTCGGAGGCTGAGGAGGAGATGGCGTGAACCCGGGAGCGGAGCTTGCAGTGAGCCA 2644
 Db CTCGGAGGCTGAGGAGGAGATGGCGTGAACCCGGGAGCGGAGCTTGCAGTGAGCGG 4972
 QY 2645 AGATTGCGCCACTGCAGTCCGCGCTGGCGGACAGAGCGAGACTCTGTTTCAAA 2704
 Db AGATTGCGCCACTGCAGTCCGCGCTGGCGGACAGAGCGAGACTCTGTTTCAAA 4912
 QY 2705 AAAAAAATAAAAAAATAAATGAGAAACCGCATCTTTTACACTTAGGTTG 2764
 Db AAAAAAATAAAAAAATAAATGAGAAACCGCATCTTTTACACTTAGGTTG 4854

```
QY 2765 AGTTTCTGTATCATATAAAGGGTTTGGATTAAAGTATGCTCCCTGGCACTTATAAATGTT 2824
DB 4853 CTTAAATGATTGCTGTATGACAGCTATTTTCATGCGGAA---TCACAAAATTTGTTATT 4797
QY 2825 AGGCTTTAATATTATATAGATCGAGTAGTATTTTCATCTTTAGTCGCTCCTTTAGTCAC 2884
DB 4796 AAAATGATTATTTAGTTGAAAAATGCACCTGCAAGATTTTTGGGATATGACACGTTAA 4737
QY 2885 TCTTCCTATACCACTGAGACCACTTTTACAATTTAGAAAAGACAAATAACTGTTGGGT 2944
DB 4736 TTAAGCGCAAACTTTAGAAAATTTATCAAAACCAAGAGGTAAATAAAGATGCC 4677
QY 2945 TACTTGATGATATAAACAACCAAGAAAATAATTTTGAAGGAATTAAGTTTGAACACACA 3004
DB 4676 TACTTAAGAAGA---AACCACATAGAATTTAGGCAAAAAGGGAAGTATGAATACAGAA 4621
QY 3005 TGTTAACAATTTTACCAAGTGGGATTTGCCTGTGATTAAGATGCTGTAAACATTTGG 3064
DB 4620 TTTAAGCCTCTATGCTAGTGTAGTATCAATAAAGTACATTATACAGATCAATAAAGTT-- 4563
QY 3065 GCCAGTAGTTATATTTGAAAAATGTTTATAGCCAATATATATTTTATTTTAAATATA 3124
DB 4562 -----CTATATGCACAGACTAATACAAAGAAATCAAGTTGAAAGAGCTCTG 4514
QY 3125 CAGTTTCATCAGTCTATAGTATTTCAATTAAGTCTAAGATGCCATCAGTGTAGCAAC 3184
DB 4513 GAGCTTGCCCACTAAGTCTCACCAGTGGG-----NA 4480
QY 3185 ACCACTGTTTATGCACTGCTAAGAAAGAAATAAAGGGCTGTGTGAGTGGCTCACACTG 3244
DB 4479 ATCCCTTCTTCAAAATGATGGCCGGCATGGTGGCTATGCTATGATCCAGCACTT 4420
QY 3245 TGGACGCGCAAGCAGAGATCACTTGAGGCCAGAAAGTTCAAGACCAACCTGTGTCAACA 3304
DB 4419 TGGAGGCGCTAGGCAGCGGATCACTTGAGGTCAGGAGTTTGAACACCGCTGGCCCAACA 4360
QY 3305 TTGTAAAGACCTGTCTACAAAAAATAAAGTTTAAATTAAGTGGTGGTGGGCGAC 3364
DB 4359 TGGTGAACCCCTCTCTACTAAATACAAACAATTAGCCAGGTGTG-TGGTGGCAC 4301
QY 3365 ATGCTGTAGTTCAGTCTACTCTGAGAGTAAAGTGGGAGGATTTCTAGAGCCAGCGGTGT 3424
DB 4300 ACGCTCTAATCCAGCTACTTGGAGGCTAAGGCAGGAGTTCAGGCTAGAAAACGGAGCAAG 4241
QY 3425 TGAAGCTGCAATGAGCTGTGACACACACCACTCCGCTCCAGCGTGGGCAACAGAGTGAGA 3484
DB 4240 TGGAGGTTGTTAAGCCGAGATTGTCCCACTGCATTCAGGCTAGAAAACGGAGCAAG 4181
QY 3485 CCCTGTTCTTAAAGAAAGAAAGAAAGGGCTGCCACCTTAACAGACACACTATTGAG 3544
DB 4180 CTCCATCTCAAAAACAAAACAGTACGTAATGAAGAGCAGCTTCCACAAAATCAAAACCTAG 4121
QY 3545 TT 3546
DB 4120 CT 4119

RESULT 9
ID AAK66446
AC AAK66446;
XX
XX
XX 06-NOV-2001 (first entry)
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:21258.
XX Human: immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
```

WO200157182-A2.

09-AUG-2001.

17-JAN-2001; 2001WO-US01354.

31-JAN-2000; 2000US-0179065.

04-FEB-2000; 2000US-0180628.

24-FEB-2000; 2000US-0184664.

02-MAR-2000; 2000US-0186350.

16-MAR-2000; 2000US-0189874.

17-MAR-2000; 2000US-0190076.

18-APR-2000; 2000US-0198123.

19-MAY-2000; 2000US-0205515.

07-JUN-2000; 2000US-0209467.

28-JUN-2000; 2000US-0214886.

30-JUN-2000; 2000US-0215135.

07-JUL-2000; 2000US-0216847.

07-JUL-2000; 2000US-0216880.

11-JUL-2000; 2000US-0217487.

11-JUL-2000; 2000US-0217496.

14-JUL-2000; 2000US-0218290.

26-JUL-2000; 2000US-0220963.

26-JUL-2000; 2000US-0220964.

14-AUG-2000; 2000US-0224518.

14-AUG-2000; 2000US-0224519.

14-AUG-2000; 2000US-0225213.

14-AUG-2000; 2000US-0225214.

14-AUG-2000; 2000US-0225266.

14-AUG-2000; 2000US-0225267.

14-AUG-2000; 2000US-0225268.

14-AUG-2000; 2000US-0225270.

14-AUG-2000; 2000US-0225447.

14-AUG-2000; 2000US-0225757.

14-AUG-2000; 2000US-0225758.

14-AUG-2000; 2000US-0225759.

18-AUG-2000; 2000US-0226279.

22-AUG-2000; 2000US-0226681.

22-AUG-2000; 2000US-0226868.

22-AUG-2000; 2000US-0227182.

23-AUG-2000; 2000US-0227009.

30-AUG-2000; 2000US-0228924.

01-SEP-2000; 2000US-0229287.

01-SEP-2000; 2000US-0229343.

01-SEP-2000; 2000US-0229344.

01-SEP-2000; 2000US-0229345.

05-SEP-2000; 2000US-0229509.

05-SEP-2000; 2000US-0229513.

06-SEP-2000; 2000US-0230437.

06-SEP-2000; 2000US-0230438.

08-SEP-2000; 2000US-0231242.

08-SEP-2000; 2000US-0231243.

08-SEP-2000; 2000US-0231244.

08-SEP-2000; 2000US-0231413.

08-SEP-2000; 2000US-0231414.

08-SEP-2000; 2000US-0232080.

08-SEP-2000; 2000US-0232081.

12-SEP-2000; 2000US-0231968.

14-SEP-2000; 2000US-0232397.

14-SEP-2000; 2000US-0232398.

14-SEP-2000; 2000US-0232399.

14-SEP-2000; 2000US-0232400.

14-SEP-2000; 2000US-0232401.

14-SEP-2000; 2000US-0233063.

14-SEP-2000; 2000US-0233064.

21-SEP-2000; 2000US-0233065.

21-SEP-2000; 2000US-0234223.

21-SEP-2000; 2000US-0234274.

25-SEP-2000; 2000US-0234997.

25-SEP-2000; 2000US-0234998.

26-SEP-2000; 2000US-0235484.

27-SEP-2000; 2000US-0235834.

27-SEP-2000; 2000US-0235836.

PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 13-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 01-DEC-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 (HUMA-) HUMAN GENOME SCI INC.
 Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-483426/52.
 XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX Disclosure; SEQ ID NO 21258; 307lpp + Sequence Listing; English.
 XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 11125 BP; 3006 A; 2279 C; 2356 G; 3484 T; 0 other;
 Query Match 7.5%; Score 324.6; DB 22; Length 11125;
 Best Local Similarity 58.7%; Pred. No. 7.6e-52;
 Matches 667; Conservative 1; Mismatches 440; Indels 29; Gaps 5;
 QY 2395 CCTAAAAATAAACCATAGGCGCGGCTGAGTACGCTGTAATCCACGACTTTG 2454
 Db 9321 CATAACAAATTTATATAGGCCAGCGCTGGTGGCTTANTCCTAGCACTTTG 9380
 QY 2455 GGAGGCGGAGCGGGTGGATCATGAGGTGAGAGATCGAGCATCTCGGTCTAACAGGT 2514
 Db 9381 AGAGGCGGAGCGGGGATCATGAGGCGGAGGATCGAGCATCTCGGTCTAACAGGT 9440
 QY 2515 GAAACCCGCTCTACTAAAAATACAAAAATTTAGCGCGCGGTGGCGGCGCTGTA 2574
 Db 9441 GAAACCTGCTCTACTAAAAATAAAAAAATTTAGCGCGCGGTGGCGGCGCTGTA 9500
 QY 2575 GTCCAGCTTCTCGGAGGCTGAGGAGAGATGCGGTGAACCCGCGGAGCGGAGCTTG 2634
 Db 9501 GGCCAGCTACTGGGAGGCTGAGGAGGAGATGGTGAATGGCGGAGCGGAGCTTG 9560
 QY 2635 CAGTGAGCCAAAGATTGCGCCACTGCGAGTCCGCGCTGGCGGAGCAGCAGACT 2694
 Db 9561 CAGTGAGCCGAAATTCGCGCACT- - - - - GCACCTCAGCCCTTGTGACAGCAGCAGACT 9613
 QY 2695 CTGTYTCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 2753
 Db 9614 CTGTCCTCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 9673
 QY 2754 CTTAGGGTTTGAGTTCTCTATCTATATAAAAGGGTTTGATTAAGTGATCCTGCACT 2813
 Db 9674 AGCCCATTTCTGGGATGCTTCATAGTAATAGATGCTCCCTATTCAGGGCTTAAATTT 9733
 QY 2814 TATAAAATGTTAGGGCTTAATATTTATCATAGATCAGGATAGTTTTCATCTTACTGCC 2873
 Db 9734 GTTCCCTAATCCAGAAATTTCTCTAGATACCAATATGGCTTATTTCTTATTCCTTC 9793
 QY 2874 TCCTTAGTCACTCTCTATACCAATCTGAGACCAATTTTACAAATTTAGAAAAGACAATA 2933
 Db 9794 TGGTTACTTTTAAATGTCACCTTATCAGAGAGGCGCTTCTTACTATTTCTCTCTGTTA 9853
 QY 2934 ACTGGTTGGTTTACTGATAGTATAATAACCAAGAAAAAATAATTTTAGAAGGAA -TTAAG 2992
 Db 9854 TACTCTAGGTGTTTATCCTGCTGTTTTCATATAATTTTGAATGTTGTTGTTGTTTCAT 9913

PR	08-NOV-2000;	2000US-0245524.
PR	08-NOV-2000;	2000US-0245525.
PR	08-NOV-2000;	2000US-0245526.
PR	08-NOV-2000;	2000US-0245527.
PR	08-NOV-2000;	2000US-0245528.
PR	08-NOV-2000;	2000US-0245532.
PR	08-NOV-2000;	2000US-0245609.
PR	08-NOV-2000;	2000US-0245610.
PR	08-NOV-2000;	2000US-0245611.
PR	08-NOV-2000;	2000US-0245613.
PR	17-NOV-2000;	2000US-0249207.
PR	17-NOV-2000;	2000US-0249208.
PR	17-NOV-2000;	2000US-0249209.
PR	17-NOV-2000;	2000US-0249210.
PR	17-NOV-2000;	2000US-0249211.
PR	17-NOV-2000;	2000US-0249212.
PR	17-NOV-2000;	2000US-0249213.
PR	17-NOV-2000;	2000US-0249214.
PR	17-NOV-2000;	2000US-0249215.
PR	17-NOV-2000;	2000US-0249216.
PR	17-NOV-2000;	2000US-0249217.
PR	17-NOV-2000;	2000US-0249218.
PR	17-NOV-2000;	2000US-0249244.
PR	17-NOV-2000;	2000US-0249245.
PR	17-NOV-2000;	2000US-0249264.
PR	17-NOV-2000;	2000US-0249265.
PR	17-NOV-2000;	2000US-0249297.
PR	17-NOV-2000;	2000US-0249299.
PR	17-NOV-2000;	2000US-0249300.
PR	01-DEC-2000;	2000US-0250391.
PR	01-DEC-2000;	2000US-0251160.
PR	05-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251988.
PR	05-DEC-2000;	2000US-0251988.
PR	06-DEC-2000;	2000US-0251479.
PR	08-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251868.
PR	08-DEC-2000;	2000US-0251869.
PR	08-DEC-2000;	2000US-0251990.
PR	11-DEC-2000;	2000US-0254097.
PR	05-JAN-2001;	2001US-0259678.
XX	(HUMA-)	HUMAN GENOME SCI INC.
XX		
PA	Rosen CA, Barash SC, Ruben SM;	
PI		
PX	WPI; 2001-541565/60.	
DD	Nucleic acids encoding 3224 human nervous system antigen polypeptides,	
DT	useful for preventing, diagnosing and/or treating nervous system	
PT	cancers and metastases -	
PP		
XX	Disclosure; SEQ ID NO 11443; 1701pp + Sequence Listing; English.	
PS		
XX	The invention relates to novel genes (ABA11004-ABA21534) and proteins	
CC	(ABA14678-ABR18001) useful for preventing, treating or ameliorating	
CC	medical conditions e.g. by protein or gene therapy. The genes are	
CC	isolated from a range of human tissues disclosed in the specification.	
CC	The nucleic acids, proteins, antibodies and (ant)agonists are useful	
CC	in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast	
CC	and ovarian cancer and other cancers of the adrenal gland, bone, bone	
CC	marrow, breast, gastrointestinal tract, liver, lung, or urogenital;	
CC	(b) immune disorders e.g. Addison's disease, allergies, autoimmune	
CC	hemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's	
CC	disease, multiple sclerosis, rheumatoid arthritis and ulcerative	
CC	colitis; (c) cardiovascular disorders such as myocardial ischaemias;	
CC	(d) wound healing; (e) neurological diseases e.g. cerebral anoxia and	
CC	epilepsy; and (f) infectious diseases such as viral, bacterial, fungal	
CC	and parasitic infections.	
CC	Note: The sequence data for this patent did not form part of the	
CC	printed specification, but was obtained in electronic format directly	
CC	from WFO at ftp.wipo.int/pub/published_pct_sequences.	

Db	11273	TAAAAA-----TAAAAAATTAGCTGAGTGTGTGGTGGCACATGCCTGTGTATCCCGAGCTAC	11221
Qy	3385	TCGGAGGCTAAGGTGGGAGGATTGCTAGAGCCACCGTGTGTGGAAGCTGCAATGAGCTGT	3444
Db	11220	TCGGAGGCTGAGATAGGAATCACTTGAACCCAGAGGTGGAGATTGCAGTGGAGCAA	11161
Qy	3445	GACCACACCACTCGCTCCAGCTGGCGCAACAGAGTGAGACCCCTGTTCTTAAAGAAAGA	3504
Db	11160	GAAGCGCCCACTGCACTCCAGCTGGCGCAACAGAGTGAGACTGTGTCTCAAAAAA	11101
Qy	3505	AAGAAAAA 3513	
Db	11100	CAAAAGCAA 11092	
RESULT 13			
ABQ87681/C			
ID	ABQ87681	standard; DNA; 465237 BP.	
XX	AC	XX	
XX	AC	XX	
XX	AC	XX	
DT	18-SEP-2002	(first entry)	
XX	Human	oestrogen receptor alpha gene.	
DE	Human	oestrogen receptor; oestrogen receptor alpha; cytosolic;	
KW	Human; oestrogen; receptor; oestrogen receptor alpha; cytosolic;		
KW	osteopathic; cardiant; cancer; osteoporosis; cardiovascular disorder;		
KW	chromosome 6q25.1; gene; ds.		
XX	XX		
OS	Homo sapiens.		
XX	XX		
FH	Key	Location/Qualifiers	
FT	variation	replace(18783,T)	
FT		/*tag= a	
FT	variation	/standard_name= "Single nucleotide polymorphism"	
FT		/*tag= a	
FT	exon	replace(18937,C)	
FT		/*tag= a	
FT		/standard_name= "Single nucleotide polymorphism"	
FT		18941..19032	
FT		/*tag= b	
FT		/number= 1	
FT	intron	19033..52817	
FT		/*tag= b	
FT		/number= 1	
FT	variation	/cons_splice= (5'site:YES,3'site:NO)	
FT		replace(19034,C)	
FT		/*tag= a	
FT	exon	/standard_name= "Single nucleotide polymorphism"	
FT		52818..52940	
FT		/*tag= b	
FT	variation	replace(52877,A)	
FT		/number= 2	
FT		/*tag= a	
FT	variation	/standard_name= "Single nucleotide polymorphism"	
FT		replace(52901,A)	
FT		/*tag= a	
FT	intron	/standard_name= "Single nucleotide polymorphism"	
FT		52941..64149	
FT		/*tag= b	
FT		/number= 2	
FT	exon	/cons_splice= (5'site:YES,3'site:NO)	
FT		64150..64280	
FT		/*tag= b	
FT	intron	/number= 3	
FT		64281..166227	
FT		/*tag= b	
FT		/number= 3	
FT	variation	/cons_splice= (5'site:YES,3'site:NO)	
FT		replace(64331,G)	
FT		/*tag= a	
FT	exon	/standard_name= "Single nucleotide polymorphism"	
FT		166228..166322	
FT		/*tag= b	

```

FT FT      /tag= b
FT FT      /number= 10
FT FT      replace(308292,A)
FT FT      /tag= a
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(308382,G)
FT FT      /tag= a
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      306504..373639
FT FT      /tag= b
FT FT      /number= 10
FT FT      373640..373778
FT FT      /tag= b
FT FT      /number= 11
FT FT      373779..422963
FT FT      /tag= b
FT FT      /number= 11
FT FT      422964..423097
FT FT      /tag= b
FT FT      /number= 12
FT FT      replace(423067,C)
FT FT      /tag= a
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      423098..456353
FT FT      /tag= b
FT FT      /number= 12
FT FT      replace(423149,G)
FT FT      /tag= a
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(423163,G)
FT FT      /tag= a
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(423220,A)
FT FT      /tag= a
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(423232,G)
FT FT      /tag= a
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(423258,G)
FT FT      /tag= a
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      456354..456537
FT FT      /tag= b
FT FT      /number= 13
FT FT      456538..460700
FT FT      /tag= b
FT FT      /number= 13
FT FT      replace(459706,C)
FT FT      /tag= a
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(459832,A)
FT FT      /tag= a
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(459913,G)
FT FT      /tag= a
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(460024,G)
FT FT      /tag= a
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(460056,T)
FT FT      /tag= a
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(460159,G)
FT FT      /tag= a
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(460553,C)
FT FT      /tag= a
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(460564,A)
FT FT      /tag= a
FT FT      /standard_name= "Single nucleotide polymorphism"

```

Query Match 7.3%; Score 317.4; DB 24; Length 465237;

```

Best Local Similarity 93.5%; Pred. No. 3.5e-50;
Matches 330; Conservative 1; Mismatches 22; Indels 0; Gaps 0;

QY 2387 TCTTTTCATCTAAATAAATAAACAATAGGCGGCGGCTGCTCAGCCTGTATATCCCA 2446
    |||||  |||  |||  |||  |||||  |||||  |||||  |||||  |||||  |||||
DB 47735 TATTTCTCTTAAGGACAAAGAGCGCAGTGGCGGCGGCTCAGCCTGTATATCCCA 47676

QY 2447 GCACCTTTGGAGGCGGCGGCTGGATCATGAGTFCAGGAGATCGAGACCATCTGGCT 2506
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 47675 GCACCTTTGGAGGCGGCGGCTGGATCATGAGTFCAGGAGATCGAGACCATCTGGCT 47616

QY 2507 AACAGGTGAACCCCTCTCTACTATAAATAAATAAATAAATAGCGGGCGGCTGGCGG 2566
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 47615 AACAGGTGAACCCCTCTCTACTATAAATAAATAAATAAATAGCGGGCGGCTGGCGG 47556

QY 2567 CGCCTGTAGTCCCAAGCTTCTCGGAGGCTGAGCAGGAGAAATGGCGTGAACCCGGGAGG 2626
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 47555 CGCCTGTAGTCCCAAGCTTCTCGGAGGCTGAGCAGGAGAAATGGCGTGAACCCGGGAGG 47496

QY 2627 GGAGCTTGCAGTGAGCAAGATTGCGCCACTGCGAGTCCGAGTCCGGGCTGGGGGACAGA 2686
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 47495 GGAGCTTGCAGTGAGCAAGATTGCGCCACTGCGAGTCCGAGTCCGGGCTGGGGGACAGA 47436

QY 2687 GCGAGACTCTCTTTCACAAAAAATAAATAAATAAATAAATAAATCAATATCAGGA 2739
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 47435 GCGAGACTCTCTTTCACAAAAAATAAATAAATAAATAAATAAATAAATAAATCAATATCAGGA 47383

RESULT 14
ABA90193/c
ID ABA90193 standard; DNA; 465237 BP.
XX AC ABA90193;
XX DT 11-FEB-2002 (first entry)
XX DE Human oestrogen receptor alpha gene.
XX KW Human: oestrogen receptor alpha; ESR-alpha; ER: chromosome 6; Syne-2;
XX KW synaptic nuclei expressed gene 2; haplotype; cytostatic; osteopathic;
XX KW cardiant; vasotropic; gene therapy; vaccine; cancer; osteoporosis;
XX KW cardiovascular disease; oestrogen receptor; ds.
XX OS Homo sapiens.
XX PN WO200162969-A2.
XX PD 30-AUG-2001.
XX PF 20-FEB-2001; 2001WO-US05358.
XX PR 22-FEB-2000; 2000US-183756P.
XX PR 20-OCT-2000; 2000US-0692414.
XX PR 24-JAN-2001; 2001US-0768184.
XX PA (PEKE ) PE CORP NY.
XX PI Kalush F, Cassel MJ, Hwang SS, Winn-Deen ES;
XX WPI; 2002-041152/05.
XX P-PSDB; AAG68251.
XX Novel variant of estrogen receptor alpha polypeptide useful for
PT determining the biological activity of a protein for high throughput
PT screening and for raising antibodies that elicit an immune response in
PT host -
XX Example; Fig 1 page 1-93; 333pp; English.
PS The present invention describes an isolated peptide (I) consisting of an
CC amino acid sequence selected from: (a) the amino acid sequence of a
CC variant of the oestrogen receptor alpha (ESR-alpha) protein in AAG68251;
CC or (b) a fragment comprising at least 10 contiguous amino acids of the

```

us-09-730-559b-7.rng

PR	07-JUN-2000;	2000US-0209467.
PR	28-JUN-2000;	2000US-0214886.
PR	30-JUN-2000;	2000US-0215135.
PR	07-JUL-2000;	2000US-0216647.
PR	07-JUL-2000;	2000US-0216680.
PR	11-JUL-2000;	2000US-0217487.
PR	11-JUL-2000;	2000US-0217496.
PR	14-JUL-2000;	2000US-0218290.
PR	26-JUL-2000;	2000US-0220963.
PR	26-JUL-2000;	2000US-0220964.
PR	14-AUG-2000;	2000US-0224518.
PR	14-AUG-2000;	2000US-0224519.
PR	14-AUG-2000;	2000US-0225213.
PR	14-AUG-2000;	2000US-0225214.
PR	14-AUG-2000;	2000US-0225266.
PR	14-AUG-2000;	2000US-0225267.
PR	14-AUG-2000;	2000US-0225268.
PR	14-AUG-2000;	2000US-0225470.
PR	14-AUG-2000;	2000US-0225477.
PR	14-AUG-2000;	2000US-0225757.
PR	14-AUG-2000;	2000US-0225758.
PR	14-AUG-2000;	2000US-0225759.
PR	18-AUG-2000;	2000US-0226681.
PR	22-AUG-2000;	2000US-0226682.
PR	22-AUG-2000;	2000US-0226686.
PR	22-AUG-2000;	2000US-0227182.
PR	23-AUG-2000;	2000US-0227909.
PR	30-AUG-2000;	2000US-0228924.
PR	01-SEP-2000;	2000US-0229287.
PR	01-SEP-2000;	2000US-0229343.
PR	01-SEP-2000;	2000US-0229344.
PR	01-SEP-2000;	2000US-0229345.
PR	05-SEP-2000;	2000US-0229509.
PR	05-SEP-2000;	2000US-0229513.
PR	06-SEP-2000;	2000US-0230437.
PR	06-SEP-2000;	2000US-0230438.
PR	08-SEP-2000;	2000US-0231242.
PR	08-SEP-2000;	2000US-0231243.
PR	08-SEP-2000;	2000US-0231244.
PR	08-SEP-2000;	2000US-0231413.
PR	08-SEP-2000;	2000US-0231414.
PR	08-SEP-2000;	2000US-0232080.
PR	08-SEP-2000;	2000US-0232081.
PR	14-SEP-2000;	2000US-0232397.
PR	14-SEP-2000;	2000US-0232398.
PR	14-SEP-2000;	2000US-0232399.
PR	14-SEP-2000;	2000US-0232400.
PR	14-SEP-2000;	2000US-0232401.
PR	14-SEP-2000;	2000US-0233063.
PR	14-SEP-2000;	2000US-0233064.
PR	14-SEP-2000;	2000US-0233065.
PR	21-SEP-2000;	2000US-0234223.
PR	21-SEP-2000;	2000US-0234274.
PR	25-SEP-2000;	2000US-0234997.
PR	25-SEP-2000;	2000US-0234998.
PR	26-SEP-2000;	2000US-0235484.
PR	27-SEP-2000;	2000US-0235834.
PR	27-SEP-2000;	2000US-0235836.
PR	29-SEP-2000;	2000US-0236327.
PR	29-SEP-2000;	2000US-0236367.
PR	29-SEP-2000;	2000US-0236368.
PR	29-SEP-2000;	2000US-0236369.
PR	29-SEP-2000;	2000US-0236370.
PR	02-OCT-2000;	2000US-0236802.
PR	02-OCT-2000;	2000US-0237037.
PR	02-OCT-2000;	2000US-0237038.
PR	02-OCT-2000;	2000US-0237039.
PR	02-OCT-2000;	2000US-0237040.
PR	13-OCT-2000;	2000US-0239935.
PR	13-OCT-2000;	2000US-0239937.
PR	20-OCT-2000;	2000US-0240960.
PR	20-OCT-2000;	2000US-0241221.

14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.

14-AUG-2000; 2000US-0225757.
14-AUG-2000; 2000US-0225758

14 -AUG-2000; 2000US -0225759.

PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-022668

PR 22-AUG-2000; 2000US-0227182.

PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287

PR 01-SEP-2000; 2000US-0229343.

PR 01-SEP-2000; 2000US-0229345.

PR 05-SEP-2000; 2000US-0229513.

PR 06-SEP-2000; 2000US-0230438.

PR 08-SEP-2000; 2000US-0231243.

PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.

PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081

PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397

PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399

PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.

PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.

PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223

PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.

PR 25-SEP-2000; 2000US-0234998.

PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.

PR 29-SEP-2000; 2000US-0236327.
20-SEP-2000; 2000US-0236367

PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369

29-SEP-2000; 2000US-0236370.
03-OCT-2000; 2000US-0236802

PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.

PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040

PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.

PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221

Db 46192 TGACATCATAGGCTGG---GCATGGTGGCTCACACCTGTAACTCTAACACTCTTGGGAGGC 46248
Qy 3253 CAAGGCGAGGAGCATCACTTTGAGGCCAGAAAGTTCAAGACCACCTGGTCAACATTGTAAAGA 3312
Db 46249 TGAGGTGGGAAGACTGCTTAAGGCCACGGGTTTGAGGCCACCTGGGCAACACTGCG-GA 46307
Qy 3313 CCCTGTCTCTACAAAAAAGTTAAAAATTAGTGGGTGGGTGGGCAACATGCCTGT 3372
Db 46308 CCCGCTCTCTACAAAAATAAAAA---AAAAATTAGCCAGGCTTGGTGGCACAATGCCTGC 46364
Qy 3373 AGTTCCAGCTACTCTGGAGGCTAAGGTGGGAGGATTGCTAGAGCCACGSGTTTGGAGCT 3432
Db 46365 AGTCCCAAGCTACTTGGGAAGCTGAGGTGGGAGGCTACCTGAGCCTGGGAGTTTGGGCT 46424
Qy 3433 GCAATGAGCTGTGACCACACCTGCCGCTCCAGCGTGGGCAACAGAGTGAGACCCCTGTTT 3492
Db 46425 GCAGTGAGCTAGGACTGTACCACCTGCACCTCCAGTCTGAGCCACAGAGCGAGACCCCTGCT 46484
Qy 3493 CTAAAGAAAGAAAGAAAGAGGCTGCCA 3522
Db 46485 CAAAAAACAAGAAAGAAAGAGATGGCA 46514

Search completed: February 28, 2003, 10:18:48
Job time : 6002 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 28, 2003, 04:00:27 ; Search time 134 Seconds
(without alignments)
9939.531 Million cell updates/sec

Title: US-09-730-559B-7
Perfect score: 4343
Sequence: 1 atggacgccccaaagcagg.....attcacagcagtcctctct 4343

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	336.4	7.7	246240	2 US-08-724-394A-20	Sequence 20, Appl
2	336.4	7.7	246240	2 US-08-724-394A-21	Sequence 21, Appl
3	336.4	7.7	246240	2 US-08-724-394A-22	Sequence 22, Appl
4	310.8	7.2	176373	3 US-09-128-155-17	Sequence 17, Appl
5	309	7.1	152331	3 US-09-128-155-16	Sequence 16, Appl
6	303.6	7.0	380	1 US-08-126-587C-5	Sequence 5, Appl
7	291.4	6.7	3609	4 US-09-705-299-11	Sequence 11, Appl
8	291.2	6.7	36159	4 US-09-749-588-3	Sequence 3, Appl
9	289.6	6.7	59065	4 US-09-813-817-3	Sequence 3, Appl
10	289.6	6.7	59065	4 US-09-978-197-3	Sequence 3, Appl
11	288.6	6.6	8342	3 US-08-545-860D-63	Sequence 63, Appl
12	288.6	6.6	8342	5 PCr-US94-04496-63	Sequence 63, Appl
13	288.6	6.6	8392	3 US-08-080-255-6	Sequence 6, Appl
14	288.6	6.6	8392	3 US-08-465-713-6	Sequence 6, Appl
15	288.6	6.6	8392	5 PCr-US93-05857-6	Sequence 6, Appl
16	286.2	6.6	5095	1 US-08-092-817-3	Sequence 6, Appl
17	286.2	6.6	5095	4 US-08-485-128-3	Sequence 3, Appl
18	283.8	6.5	162450	4 US-09-345-882-1	Sequence 1, Appl
19	280.8	6.5	4220	1 US-08-832-883-66	Sequence 66, Appl
20	280.8	6.5	4220	2 US-08-832-877-66	Sequence 66, Appl
21	280.6	6.5	84495	4 US-09-797-906-3	Sequence 66, Appl
22	278.2	6.4	15977	4 US-09-608-285A-59	Sequence 59, Appl
23	276.8	6.4	14747	4 US-09-608-285A-42	Sequence 42, Appl
24	273	6.3	21234	4 US-09-810-671-3	Sequence 3, Appl
25	271.8	6.3	14636	4 US-09-173-914-6	Sequence 6, Appl
26	271.8	6.3	26664	4 US-09-564-803-28	Sequence 28, Appl
27	267.6	6.2	1751	4 US-09-149-476-110	Sequence 110, App

ALIGNMENTS

RESULT 1
US-08-724-394A-20
; Sequence 20, Application US/08724394A
; Patent No. 587237
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Krommal, Gregory S.
; APPLICANT: Lauer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 587237el
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/724, 394A
; APPLICATION NUMBER: US/08/724, 394A
; FILING DATE: 01-Oct-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 017957-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 246240 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..246240
; OTHER INFORMATION: /note= "HLA-H.CONTIG"
US-08-724-394A-20

Sequence 5, Appli
Sequence 10, Appl
Sequence 10, Appl
Sequence 5, Appli
Sequence 11, Appl
Sequence 11, Appl
Sequence 11, Appl
Sequence 5, Appli
Sequence 1, Appli
Sequence 3, Appli
Sequence 10, Appl
Sequence 20, Appl
Sequence 20, Appl
Sequence 20, Appl
Sequence 20, Appl
Sequence 20, Appl

Qy	2746	TC	TTTACAC	TTAGG	TTTGAG	TTTCTG	TATCT	ATATA	AAAGG	TTTGG	ATTAAG	TGATCC	2805													
Db	107408	TG	CA	TTTGC	TTTGGG	ACTTTTA	TATAT	TAGTCT	TACAA	ATCTAG	CCACCA	TAGAAATCTGC	107467													
Qy	2806	CT	GGC	ACTTTA	AAATCT	TAGG	CTTA	ATA	TATAT	TCAT	AGATCG	AGGATAG	TTTCATCTCT	2865												
Db	107468	TG	ATTA	ATACGG	TTCTGT	TAA	ATGGA	ACA	TGCA	TTTGGG	GAAAA	AGAGGG	AGT	107527												
Qy	2866	TAG	TG	CGCTC	CTTAG	TCA	CTCT	CC	TATAC	CAATCTG	AGACCA	TTTACA	ATTTAG	2925												
Db	107528	GT	TTT	TAG	TGAT	TTTGT	TTT	TAC	TTG	TTTAA	TAAATTT	TAA	AGCAATCT	107587												
Qy	2926	GAC	AA	TAA	CTGG	TTGG	TTTAC	T	TGAT	AGT	ATA	TAAT	TAACCA	AGNAAT	TA	TTT	TAG	AGG	2985							
Db	107588	A	CA	TTTTAT	TTCT	AC	TG	TG	AACT	TGCA	TAA	GGT	TAT	GA	GAT	AA	AGT	T	TACA	GC	TAT	TCA	107647			
Qy	2986	AA	TTA	AGTTG	GA	AA	CCAC	ATG	-	TTA	CA	AAAT	TTCT	AC	CA	AA	GGTGG	-	GA	TTT	GC	GT	GAT	3042		
Db	107648	CA	TAC	AG	TTTGT	AG	CTTT	A	TA	AT	TAT	GA	AT	TCT	AC	AG	AT	TA	AT	T	GC	T	TA	TAT	GAT	107707
Qy	3043	TAA	AG	TG	CT	TA	AA	CA	NTTGG	CG	CA	TAG	TTA	TA	AT	TTG	AA	AA	TG	TTT	TAG	CA	CA	TA	3102	
Db	107708	GAA	AT	CT	CA	TA	AT	ACA	TTAG	CA	TAT	AT	TTT	TA	TAA	CA	CA	AT	T	CAG	AA	GAG	AG	CA	TA	107767
Qy	3103	TAT	AA	TTTTTT	TAT	TTAA	AT	ATAC	AG	TTTTC	AT	CA	TAG	CT	CTA	-	----	-	TT	AG	TAT	TT	CA	TAA	3155	
Db	107768	CC	CA	NTT	TC	AA	AT	CA	TATTA	TTG	TGA	AA	TAT	TTAG	GC	AG	CC	AAAA	TAT	TT	CT	GGA	AT	107827		
Qy	3156	GT	CT	TA	AG	AT	GC	CA	TG	TTT	AG	CA	AA	CC	CA	CT	CTTT	TAT	GC	AT	CT	GA	AA	AG		

RESULT 4

```

US-09-128-155-17
; Sequence 17, Application US/09128155
; Patent No. 6117654
;
; GENERAL INFORMATION:
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: 09404/052001
; CURRENT APPLICATION NUMBER: US/09/128,155
; CURRENT FILING DATE: 1998-08-03
; EARLIER APPLICATION NUMBER: US 60/091,650
; EARLIER FILING DATE: 1998-07-02
; EARLIER APPLICATION NUMBER: US 60/054,646
; EARLIER FILING DATE: 1997-08-04
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 176373

```

:	TYPE: DNA
:	ORGANISM: Homo sapiens
:	FEATURE:
:	NAME/KEY: misc_feature
:	LOCATION: (1)...(176373)
:	OTHER INFORMATION: n = A,T,C or G
US -09-128-155-17	
Query Match	7.2%; Score 310.8; DB 3; Length 176373;
Best Local Similarity	62.4%; Pred. No. 3.3e-59;
Matches 703; Conservative 1; Mismatches 348; Indels 74; Gaps	
Qy 2397 TAAAAATAAAACCATAGCGCGGCGGGTGGCTCAGCCCTGTAAATCCAGCACTTTGGG 2456	
Db 64735 TTAAAAATGCCACTCTTGCGCGGCGGGTGGCTCAGCCCTGTAAATCCAGCACTTTGGG 64794	
Qy 2457 AGCCCGAGCGGGTGATCATGAGTCAGGAGATCGAGAGACATTCCTGGCTACAAGGTGA 2516	
Db 64795 AGCCCGAGCGGGTGATCATGAGTCAGGAGATCGAGAGACATTCCTGGCTACAAGGTGA 64854	
Qy 2517 AACCCCGTCTCTACTAAAAATACAAAAAATTAGCCGGCGCGGTGCGCGGCCTGTAGT 2576	
Db 64855 AACCCCGTCTCTACTAAAAATACAAAAAATTAGCCGGCGCGGTGCGCGGCCTGTAGT 64914	
Qy 2577 CCCAGCTCTCCGGAGGCTGAGGCAGGAGATGCGCTGAACCCGGAGCGGAGCTTCCA 2636	
Db 64915 CCAGAGTACTCGGAGGCTGAGGCAGGAGATGCGCTGAACCCGGAGCGGAGCTTCCA 64974	
Qy 2637 GTGAGCCAAGATTGGCCCACTGCAGTCCGCGAGTCCGGCTTGGCGCACAGAGCACTCT 2696	
Db 64975 GTGAGCCGAGATTGGCCCACTGCAGTCCGCGAGTCCGGCTTGGCGCACAGAGCACTCC 65034	
Qy 2697 GTTCAAAAAAATAAAAAAAAAAAAAAACCAATAATGAGGAACGCATCTTTACACTT 2756	
Db 65035 GTCTCAAAAAAATAAAAAAAAAAAAAATGCCACTCTTGTGTA --- AAATGACCCCTG 65091	
Qy 2757 AGGTTTGTAGTTTCTGTATCTATAAAAGGGTTTTGGATTAAAGTGATCCCTGGCACTAT 2816	
Db 65092 GGCGAGGAGGAGTAGAATGTCAAAGTGAAGCACACCACCTCAGGAGTCAAGGCAAT 65151	
Qy 2817 AAAATGTTAGGCTTAATATTATTCATAGATCGAGGATAGTTTCATTCTTTAGTCGCTCC 2876	
Db 65152 GGACTGTGCAGAGAGACTGACATCTTAGACTCGGGCAA -----TAGAGAGAAG 65201	
Qy 2877 TTAGTCACCTTTCCATACCAATCTGAGACCATTTTACAATTTAGAAAAAGACAAATACT 2936	
Db 65202 GTCGTGAGGATT-----ATATTCTGGGCATAAAGCAACAGAACTAGCT 65245	
Qy 2937 GGTGTGGTTTACTTGATAGTATAATAACCAAGAAAAATAATTTTAGAGGAATAATTAGTTTG 2996	
Db 65246 GATGGCGT-----CAACCTAGAGATCAGGGAAGAAAGAAATCAAGGG 65290	
Qy 2997 AAACCAATGTTTAAACAAATCTTACAAAGTGGGATTTGCCCTGTGATTTAAGATGCTGTA 3056	
Db 65291 CATTCATAGGTTTGAGGGTTGAGTAACCTGGGGAT-----TTACAGANAATGGAGAAG 65344	
Qy 3057 ACATTTGGCCAGTAGTTATTAATTTGAAAAATGTTTATAGCCAATATAAATTTTTTAAT 3116	
Db 65345 TCTGGGAAGGGCAAGTATCTGGGGCAGGGCTCAAAAGTTCTGTATTTTGGCCAAGT 65404	
Qy 3117 TAAATATACAGTTTCATCAGCTATTAGTATTTCATTTAAGTCTAAGATGCCATCAGTGT 3176	
Db 65405 TAATTAATATTTGATACCTCT-----TAGTGTCGAAGTGAAGATGTCAACACAGTCAA 65459	
Qy 3177 TAGCAACACCACTGTTTTATGCATGCTAAGAAGAAATAAAGGGCTGTGTGCAGTGGCT 3236	
Db 65460 TTGAATACAAA---TCTGAATCTTAGCCCAGGATGGTCTCACACCTGT-----AATCC 65510	
Qy 3237 CAGACCTGTGGAGCGCAAGGCAGAGACATCCTTTAGGCCAGCAAGTTCAAGACCAACT 3296	
Db 65511 CAGCACTTTGGAGGCTGAGGTGAGAGATCACTTTGGAGCCAGGATTTGTGTATCAGCT 65570	
Qy 3297 GGTCAACATTTGAAGACCTGTCTCT--TACAAAAAATAAAGTTAAAAATTTAGCTGGCT 3354	

Qy	3019	ACCAAAAGTGGGATTTGGCTGTGATTTAAAGATGCTGTTAAACATTTGGCGCCAGTAGTTATAA	3078
Db	15244	CCTCA-----CAGTATTACTGTAATAAAGTGTGTAGCAGCTGGGAACACTATTCCAGAG	15299
Qy	3079	TTTGA AAAATGTTTATAGCCAAATATATAATTTTTTTATTAAATATACAGCTTTCATCAGTC	3138
Db	15300	AGCACTCATGAATGTTTGTCTTTGTTATTAGTTACTAGAGGCCAAATGCTGCCAGGG	15359
Qy	3139	TATTAGTATTTTCATTAAAGTCTAAGATGCCATCAGTGGTTAGCAAAACACCACTGTTTATG	3198
Db	15360	CTGAATAAATATGTGCAATTTGGTGATTTGCCACATATCTAAAGAAGTAGTTATTTTTTTT	15419
Qy	3199	CAGTCG-----TAAGAAAGAAATAAAGGGCTGTGTGCAGTGGCTCACACCTG-----	3244
Db	15420	CAATTAAACTTTAGTTTAAAAACCANTATAGGCCGCGCAGTGGCTCACACTGTAT	15479
Qy	3245	-----TGGGACGCCAAGCGAGGAGCATCACTTTAGGCCGAGAAGTTTCAAGACCAAC	3294
Db	15480	CCGAGCACCTTTGGGAGGCGGAGTGGGCGAGATCATTTGAGGTCCAGGACTTCGAGACTAGC	15539
Qy	3295	CTGGTCAACATTTGAAGACCTGTCTCTACAAAAAATAAAGTT-----AAAAATTAGC	3349
Db	15540	CTGGCCACATCGTGGAAACCCCTGTCTGTCTTAAAAAAAATAAAGTACAAAAATTAGC	15599
Qy	3350	TGGGTCGGTGGGCACATGCCTGTGTAGTTCCAGCTACTCTGGAGGCTTAAGTGGGAGGATG	3409
Db	15600	CAGGCATGATGGCAGGTCCCTGTATCCAGCTACTTTGGGAGGCCGAGCAGGAATTTG	15659
Qy	3410	CTAGAGCCACGGTGTGGAAGCTGCAATGAGCTGTGACCACACCTGCGCTCCAGCGTGG	3469
Db	15660	CTTGAACCCAGAGGTGGAGGTTGTAGTGAGCCGAGTTTGTGCCACTGCACCTTCAGCCTG	15719
Qy	3470	GGCAACAGAGTGAGACCCCTGTTTCTTAAAGAAAGAAAGAAAA	3511
Db	15720	GGTACAGAGGAGACACTGTCTCAAAAAAATAAAGAAAAA	15761

```

RESULT 10
US-09-978-197-3
; Sequence 3, Application US/09978197
; Patent No. 6403353
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001178DIV
; CURRENT APPLICATION NUMBER: US/09/978,197
; CURRENT FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: 09/813,817
; PRIOR FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 59065
; TYPE: DNA
; ORGANISM: Human
; US-09-978-197-3

```

	Query Match	6.7%	Score 289.6;	DB 4;	Length 59065;
	Best Local Similarity	58.6%;	Pred. No. 1.2e-54;		
	Matches 693;	Conservative	1;	Mismatches 410;	Indels 78; Gaps 8;
Qy	2397	TAAAAATAAAACCATAGGCCGCGCGGTGGCTCAGCCCTGTATATCCACGACATTGGG	2456		
Db	14591	TAAAAATAATATGGAGCCAGCAGCCGCTGGCTCAGCATGTATCCACGACATTGGG	14650		
Qy	2457	AGGCCGAGCGGGTGGATCATAGCTCAGGAGATCGAGACCACATCTGGCTAAACAAGTGA	2516		
Db	14651	AGGCCGAGGGGGGATACAAAGTCAAGAGATCGAGACCATCTGGCTATACACAGTGA	14710		
Qy	2517	AACCCCGTCTCTACTAAAAATACAAAAAATATAGCCGGCGGGTGGGGGGCGCTAGT	2576		

[illegible]

RESULT 11

US-08-545-860D-63
 ; Sequence 63, Application us/08545860D
 ; Patent No. 6040140
 ; GENERAL INFORMATION:
 ; APPLICANT: Croce, Carlo
 ; APPLICANT: Canaani, Eli
 ; TITLE OF INVENTION: Diagnostics, Therapeutics and Methods
 ; TITLE OF INVENTION: for Detection and Treatment of Acute Leukemias
 ; TITLE OF INVENTION: Resulting from Chromosome Abnormalities in the All-1 Region
 ; NUMBER OF SEQUENCES: 94
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &
 ; ADDRESSEE: No. 6040140ris
 ; STREET: One Liberty Place, 46th floor
 ; CITY: Philadelphia
 ; STATE: Pennsylvania
 ; COUNTRY: USA
 ; ZIP: 19103

COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/545,860D
 ; FILING DATE: 07-MAR-1996
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US94/04496
 ; FILING DATE: 22-APR-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US92/10930
 ; FILING DATE: 09-DEC-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/327,392
 ; FILING DATE: 19-OCT-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/320,559
 ; FILING DATE: 11-OCT-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/062,443
 ; FILING DATE: 14-MAY-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/971,094
 ; FILING DATE: 30-OCT-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/888,839
 ; FILING DATE: 27-MAY-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/805,093
 ; FILING DATE: 11-DEC-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Deluca Esq., Mark
 ; REGISTRATION NUMBER: 33,229
 ; REFERENCE/DOCKET NUMBER: TJU-1262
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (215) 568-3100
 ; TELEFAX: (215) 568-3439
 ; INFORMATION FOR SEQ ID NO: 63:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 8342 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 2..265
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 595..666

```

; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2353..2484
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3032..3145
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 6788..6934
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 7967..8062
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 8304..8342
; US-08-545-860D-63

Query Match      6.6%; Score 288.6; DB 3; Length 8342;
Best Local Similarity 57.8%; Pred. No. 9.8e-55;
Matches 648; Conservative 1; Mismatches 430; Indels 42; Gaps 6;

QY 2415 GCGGGCGGGTGGCTCAGCCCTGTAATCCAGCAGCTTTGGGAGGCCGAGCGGGTGGAT 2474
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3974 GCTGGGCGACGGTGGCTCAGCCCTGTAATCCAGCAGCTTTGGGAGGCCGAGCGGGAT 4033
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2475 CATGAGGTGAGGATCGAGACCATCTGGCTAAACAGGTGAACCCCGCTCTCTACTAAA 2534
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4034 CACAAGGTCAGGAGATCCAGACCATCTGGCTAAACAGGTGAACCCCGCTCTCTATAAA 4093
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2535 AATAC-AAAAAATTAGCGGGCGGGTGGCGGGCGCTGTAGTCCAGCTTCTCGGGAGG 2593
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4094 AATACAAAAAATTAGCGGGCGGGTGGCGGGCGCTGTAGTCCAGCTTCTCGAGAGG 4153
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2594 CTGAGCAGGAGAATGGCGTGAACCCGGGAGCGAGCTTTGACGTGAGCAAGATTGGCG 2653
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4154 CTGAGCAGGAGAATGGCGTGAACCCGGGAGCGAGCTTTGACGTGAGCAAGATTGGCG 4213
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2654 CACTGCAGTCCCGCAGTCCGGCTGGCGGACAGCAGGAGACTCTGHTCAAAAAAATAA 2713
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4214 CACT-----GCCTCCGGCTGGCGGACAGCAAGCAAGTCCCTCTCAAAAAAATAA 4266
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2714 AAAAAAATAAATGAGGAACGCATCTTTACACTTAGGTTTGGAGTTTCTGT 2773
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4267 AAGTGTAGGCATATATATGCAATATGAAGTCTCTAA-----GATATATAT 4316
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2774 ATCTATAAAAGGGTTGGATTAAAGTATGATCCCTGGCACTTATAAAATTTAGGCTTAA 2833
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4317 ATTAATATTATTAGTTGGTCCAAAAGTAATTGCAGTAATAACATGGAAGATGCCATG 4376
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2834 TATTATTCATAGTCGAGGATAGTTTCATCTTAGTCGCCCTCTTAGTCACCTCTTCCTAT 2893
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4377 ACATATCTAGTGAAGAGCAGGTTACAAGATAATATAAAGCAATCCCATCTT 4436
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2894 ACCAATCTGAGACCATTTTACAATTTAGAAAAGACAAATAACTGGTTGGTTACTTGATA 2953
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4437 AGTTTGGAAAAGTGTTTTAAAGTATATCTAGAAAACATCTGGAAGATTACACCA 4496
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2954 GTATAATACCAAGAAAAATAATTTAGAGGAATTAAGTTTGAACCAACATGTTAACAA 3013
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4497 AAATATAAGAGTGTGGTTGATTATGGGTGACCTTTATTTGTTCTCTGTTTTTTTT 4556
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3014 ATTCTACCAAGTGGGATTTGCTCTGATTAAGATGCTGTAAACATTTGGGCCAGTAGT 3073
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4557 TTTTAACTCTTCTGAGTTTTTTGGAGTATGTACCACCTTTTACAATGAGGAAGAAAAAT 4616
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3074 TATAATTTGAAAAATGTTTATAGCCAATATATAATTTTTTTTTTAAATATACAGTTTTCAT 3133
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4617 AGCACAATTTTAAATAGGAAGCAGTAGTTTGTCAATTTATAGGGACATATCTCTACATCT 4676
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3134 CAGTCTATAGTATTTCAITTAAGTCAAGATGCCATCAAGTGGTTAGCAAAACACCACCTGT 3193
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4677 TTACAGTTCTTAAATTCCTGGCAGATACCTCTTGGCTTATTACTTTACCACATAAGATAT 4736
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Qy	3194	TTATGACGTGCTAAGAAAGAATAAAGGGCT-GTGTGCAGTGGCTCACACCTG-----	3244
Db	4737	GTATTCAAAAGGTGGTAAAGAAAATCCAGCTCGGTGTCAGTGGCTCAGCGCTGTAAATCCCA	4796
Qy	3245	-----TGGAGCGCCAAAGCAGGAGCATCACTTGAGGCCACGAAGTTCAAGACCAACCTGG	3298
Db	4797	GTACTTTGGAGGCTGACCCAGGAGCGCTTGAGCTCAGGAGTTCAAGACCAACGCTGA	4856
Qy	3299	TCAACATTTGAGACCTCTCTCTCAAAAAAAGTTTAAAAATTAGCTGGGTGGG	3358
Db	4857	GCACCATAGTGAGACCTCATCTCTACTAAAAAAAATTAATACCAGGCATGTAGTCAT	4916
Qy	3359	TGGCACATGC-----CTGTAGTTCCAGCTACTCTGGAGGCTTAAGTTGGGAGGATTG	3409
Db	4917	GTGCGCTGTAGTCCACGCTACTCTGTCCAGCTACTCTGGGAGGCTGAGGTGAGAGGATCA	4976
Qy	3410	CTAGAGCCACGGTGTGGAAAGCTGCANTGAGCTGTGACCCACACCACTCCGCTCCAGGCTG	3469
Db	4977	CTTGGCCCAAGAGATCGAGGCTGCAAGTGAGGCATTATCAGGCCACTGCATCTCCAGCGCTG	5036
Qy	3470	GGCAACAGAGTGAGAGCCCTGTTTCTTAAAGAAAGAAAGAAA	3510
Db	5037	GGCAACTAAGCAAGAGCCCTGTCTCAAAAAATTTTAAAAAAA	5077

```

RESULT 12
PCT-US94-04496-63
; Sequence 63, Application PC/TUS9404496
; GENERAL INFORMATION:
; APPLICANT: Croce, Carlo
; APPLICANT: Canani, Eli
; TITLE OF INVENTION: Diagnostics, Therapeutics and Methods
; TITLE OF INVENTION: for Detection and Treatment of Acute Leukemias
; TITLE OF INVENTION: Resulting from Chromosome Abnormalities in the
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &
; ADDRESS: Norris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/04496

```

```

1 CLASSIFICATION:
2
3 ATTORNEY/AGENT INFORMATION:
4
5 NAME: DeLuca Esq., Mark
6 REGISTRATION NUMBER: 33,229
7 REFERENCE/DOCKET NUMBER: TJU-1242
8
9 TELECOMMUNICATION INFORMATION:
10
11 TELEPHONE: (215) 568-3100
12 TELEFAX: (215) 568-3439
13
14 INFORMATION FOR SEQ ID NO: 63:
15
16 SEQUENCE CHARACTERISTICS:
17
18 LENGTH: 8342 base pairs
19 TYPE: nucleic acid
20 STRANDEDNESS: single
21
22 TOPOLOGY: linear
23
24 MOLECULE TYPE: cDNA
25
26 HYPOTHETICAL: NO
27
28 ANTI-SENSE: NO
29
30 FEATURE:
31
32 NAME/KEY: CDS
33 LOCATION: 2..265
34
35 FEATURE:
36
37 NAME/KEY: CDS
38

```

;	LOCATION:	595..666
;	FEATURE:	
;	NAME/KEY:	CDS
;	LOCATION:	2353..2484
;	FEATURE:	
;	NAME/KEY:	CDS
;	LOCATION:	3032..3145
;	FEATURE:	
;	NAME/KEY:	CDS
;	LOCATION:	6788..6934
;	FEATURE:	
;	NAME/KEY:	CDS
;	LOCATION:	7967..8062
;	FEATURE:	
;	NAME/KEY:	CDS
;	LOCATION:	8304..8342
;	FEATURE:	
PCT-US94-04496-63		

Query Match

Query Match	0.00	Score	330.00	Length	300
Best Local Similarity	57.8%	Pred. NO.	9.8e-55		
Matches 648: Conservative		1: Mismatches	430:	Indels	42: Gaps

QY	2415	GC	CGGCGCGGGTGGCTCAGCCCTGTAATCCGACGACTTTGGGAGGCGGAGCGCGGTGGAT	2474
Db	3974	GC	TGGGCACGGTGGCTCAGCCCTGTAATCCGACGACTTTGGGAGGCGGCAAGACGACGGCAT	4033
QY	2475	CAT	GAGGTGAGGATCGAGACCATCTGGCTACAAGGTCGAAACCCGCTCTCTACTATAA	2534
Db	4034	CAC	AGGTGAGGATCCAGACCATCTGGCTTAACACAGTGAACCCGCTCTCTATTATAA	4093
QY	2535	AAT	ATAC-AAAAAATTAGCCGGGCGGTGGCGGGCCCTGTAGTCCCGAGCTTCTCGGAGG	2593
Db	4094	AAT	ACAAAAAATTAGCCGGGTGTGGTGGCGGGCCCTGTAGTCCCGAGCTCTCGAGAGG	4153
QY	2594	CT	GAGCAGGAGGAATGGCGTGAACCCGGAGCGGAGCTTGCAGTGAGCCGAAGATTGCGC	2653
Db	4154	CT	GAGCAGGAGGAACGGCATGAACCCGGAGGTGGAGCTTGCAGTGAGCGAGATCGCC	4213
QY	2654	CAC	TGCACTCCGAGTCCGGCTGGCGGCACAGAGCGAGACTCTCTYTCAAAAAATAA	2713
Db	4214	CAC	T-----GCAC'TCCGCGCTGGCGGCACAGAGCAAGCTCCGCTCAAAAAATAA	4266
QY	2714	AAAA	AAAAAACCATAAATGAGGAACGCACTCTTACACTTAGGGTTTGAGTTCTCTGT	2773
Db	4267	AGT	GTAGGCCATATATAATGCCAATATGAAGTCTTAA-----GATAATATAT	4316
QY	2774	AT	CTATAAAAAAGGTTTGGATTAAGTGATCCCTGGCACTTATAAAATGTTAGGCTTAA	2833
Db	4317	ATT	AATATTATTAGGTTGGTCCAAAGTAATTGCAGTAATAACATGGAAGATGTCCATG	4376
QY	2834	TAT	TATTCATAGATCCAGGATAGTTTCATCTTCTAGTCGCCCTCTAGTCACTCTTCTAT	2893
Db	4377	AC	ATATCACTGAGTGAAGAACAGAGGTACAAGATAATATAAGACAAATCCCATCTT	4436
QY	2894	AC	CAATCTGAGACCAATTTTACAATTTAGAAAAAGACAAATAACTGGTTGGTTACTTGATA	2953
Db	4437	AG	TTCGAAAAGTGTTTTTAAAGTATATCTGAAAAACATCTGGAAGGATTCACACCA	4496
QY	2954	GT	ATAATAACCAAGAAAAATAATTTTAGAAGGAATTAAGTTTGAACACCATGTTTACAA	3013
Db	4497	AA	ATATTAGAGTGTGGTTGGATTTATGGGTGACCTTATTTTGTCTCTGGTTTTTTTTT	4556
QY	3014	ATT	CTACCAAGTGGGATTTGCCCTGTGATTAAAGATGCTGTAAACATTTGGCGCAGTAGT	3073
Db	4557	TTT	TAATCTTCTGAGTTTTTTTGGAGTAGTGTACCACCTTTTACAATGAGGAAGGAAAAAGT	4616
QY	3074	TAT	AATTTGAAAAATTTTATAGCCAAATATATAATTTTTTTTATTTAAATATACAGTTTCAT	3133
Db	4617	AG	CACAATTTTAAATAGGAAGCAGTAGTTTGTCATTTATAGGGACATATCTCATACCT	4676
QY	3134	CAG	TCTATTAGTATTTTATTAAAGTCTAAGATGCCATCAGTGGTTAGCAACACCACTGTT	3193
Db	4677	TT	ACAGTCTTAATTTCCCTGGCAGATACCTCTTTGGCTTATTACTTACCACATTAAGATAT	4736


```

; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/05857
; FILING DATE: 19930617
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/900,689
; FILING DATE: 17/06/92
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: ARCD:072/PAR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 320-7200
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8392 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; PCT-US93-05857-6

```

```

Query Match          6.68; Score 288.6; DB 5; Length 8392;
Best Local Similarity 57.8; Pred. No. 9.8e-55;
Matches 648; Conservative 1; Mismatches 430; Indels 42; Gaps 6;

QY 2415 GCCTGGCGCGGTGCTACGCTGTAATCCAGCACATTTGGAGGCGGAGCGGGTGGAT 2474
DB 4025 GCTGGGACGGTGGCTACGCTGTAATCCAGCACATTTGGAGGCGGAGCGGCGGAT 4084

QY 2475 CATGAGGTGAGGAGATCGAGACCATCTCTGCTAAACAGGTGAACCCCGTCTCTACTAAA 2534
DB 4085 CACAAGGTGAGGAGATCCAGACCATCTCTGCTAAACAGGTGAACCCCGTCTCTACTAAA 4144

QY 2535 AATAC-AAAAAATTAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2593
DB 4145 AATACAAAAAATTAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4204

QY 2594 CTGAGGCGAGGAGATGGCGTGAACCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 2653
DB 4205 CTGAGGCGAGGAGATGGCGTGAACCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 4264

QY 2654 CACTGCGAGTCCGCGAGTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2713
DB 4265 CACT-----GCACCTCCCGCGTGGCTTACAGCAAGACTCCGCTCTCAAAAAAATAA 4317

QY 2714 AAAAAAATAAACCAATAATAGCAACGATCTTTACACTTACGCTTACGCTTACGCTTCTGT 2773
DB 4318 AAGTGTAGGCGCATATATAATGGCAAAATATGAAGTCTCTCTCAAAAAAATAA 4367

QY 2774 ATCTATAAAAAAGGTTGGATTAAAGTCAATCCCTGGCACTATAAATAATGTAGGCGCTTAA 2833
DB 4368 ATTAATATTATAGTTGGTGCACAAAGTAATTGCGAGTAATACATGGAAGAGATGCCATG 4427

QY 2834 TATTATTCATAGATCGAGGATAGTTTCAATCTTAGTCGCTCTCTTACCTCTTCTCCAT 2893
DB 4428 ACATATCACTCAGTGAAAAAGCAGGTTACAAGATAATATAAAGCAACATCCATCTT 4487

QY 2894 ACCAATCTGACACCAATTTTACAATTTAGAAAGACAAATAACTGGTTGGGTTACTTGATA 2953
DB 4488 AGTTGGAAAAAGTGTTTTAAAGTATATATCTAGAAAAACAATCTGGAAGGATTACACACCA 4547

QY 2954 GTATAATAACCAACAAAAATAATTTAGAGGAATTAAGTTTGAACCAACATGTTAACAA 3013
DB 4548 AATATTAAAGAGTGGGTGGATTATGGTGACCTTTATTTGTTTCTCTGCTGTTTTTTTT 4607

```

```

QY 3014 ATTCTACCAAAGTGGATTGCTGATTAAAGATGCTGTAAACATTTGGGCGCACTAGT 3073
DB 4608 TTTTAACTTTCTGAGTTTTCGAGTATGTAACCACTTTTACATAGGAAGAAAAAGT 4667

QY 3074 TATAATTTGAAAAATGTTTATAGCCAAATATATAATTTTATTTAATAATATACAGTTTCAT 3133
DB 4668 AGCACAATTTTAAATAGGAAGCAGTAGTTTGTCAATTTATAAGGACATATCTCTACATCTCT 4727

QY 3134 CAGTCTATTAGTATTTCATTAAGTCTAAGATGCCATCAGTGGTTAGCAACACCACTGTT 3193
DB 4728 TTACAGTCTTAAATTCCTGGCAGATACCTCTTTTGGCTTATTACTTTACCACATAAGATAT 4787

QY 3194 TTATGCACTGCTAAGAAAGAAATAAAGGCT-GTGTGCACTGGCTCACACTG----- 3244
DB 4788 GTATTCAAAGGTGGTAAAGAAATCCACGTCGGGTGCGTCAAGCTCTGTAATCCCA 4847

QY 3245 -----TGGACCGCCCAAGGCGAGGAGCATCACCTTGGAGCCAGAGTTTCAAGACCAACTGG 3298
DB 4848 GTACTTTGGGAGGCTGACGCGAGGAGGCGCTTGAGCTCAGGAGTTTCAAGACCAAGCTGA 4907

QY 3299 TCAACATTTGTAAGACCTCTCTCTACAAAAAATAAAGTTAAAAATTAGCTGGTGGCGG 3358
DB 4908 GCACCATAGTGAGACCTCTATCTCTACTAAAAAATAAATAACACAGGATGGTAGCAT 4967

QY 3359 TGGCACATGC-----CTGTAGTTCAGCTACTCTGAGGCTAAGGTGGGAGGATTTG 3409
DB 4968 GTGCTGTAGTCCCGAGCTACTCTAGTCCAGCTACTTTGGAGGCTGAGGTGAGAGATCA 5027

QY 3410 CTAGAGCCAGGTTGTTGGAAGCTGCAATGAGCTGTGACCAACCACTGCGCTCCAGCGTG 3469
DB 5028 CTTGAGCCAGGAGATCGAGGCTGCAAGTGTGAGCTTATCAGGCCACTGCACCTCCAGCCTG 5087

QY 3470 GGCAACAGAGTGAGACCCCTGTTTCTTAAAGCAAGAAAGAAA 3510
DB 5088 GGCAACTAAGCAAGACCCCTGCTCTCAAAAAAATTTTAAAAAA 5128

```

Search completed: February 28, 2003, 09:48:00
Job time : 3326 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 28, 2003, 09:06:17 ; Search time 268 Seconds
(without alignments)
10105.882 Million cell updates/sec

Title: US-09-730-559B-7

Perfect score: 4343

Sequence: 1 atggagcccccacaaagcagg.....attcacagcagtcctccctct 4343

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 460893 seqs, 311809382 residues

Total number of hits satisfying chosen parameters: 921786

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_NA.*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	990	22.8	2867	9	US-09-924-396B-17
2	877.2	20.2	3770	9	US-09-924-396B-19
3	345.8	8.0	16511	10	US-09-764-869-2064
4	317.4	7.3	465237	10	US-09-933-267A-1
5	311.8	7.2	368004	10	US-09-949-654-3
6	311	7.2	20210	10	US-09-764-870-598
7	310.8	7.2	176373	9	US-10-095-407-17
8	310.2	7.1	172637	10	US-09-805-458A-3
9	309.6	7.1	13873	9	US-09-764-868-1282
10	309.2	7.1	27007	9	US-09-764-868-1424
11	309	7.1	152331	9	US-10-095-407-16
12	307	7.1	110096	10	US-09-880-107-1542
13	306.4	7.1	2275	10	US-09-764-864-209
14	305.6	7.0	31730	10	US-09-764-877-3810
15	305.4	7.0	326014	10	US-09-731-231A-3
16	305.2	7.0	1503841	9	US-09-946-807-1
17	305.2	7.0	1503841	10	US-09-795-668-1
18	305.2	7.0	1503841	10	US-09-795-686-1
19	304.2	7.0	180557	12	US-10-003-806-6

c	20	304.2	7.0	180557	12	US-10-003-806-9	Sequence 9, Appli
c	21	302.6	7.0	4206	10	US-09-764-869-1884	Sequence 1884, Ap
c	22	301.6	6.9	327	10	US-09-764-877-2571	Sequence 2571, Ap
c	23	301.4	6.9	15643	10	US-09-764-870-638	Sequence 638, App
c	24	300.4	6.9	31314	10	US-09-764-877-3875	Sequence 3875, Ap
c	25	299.2	6.9	18000	10	US-09-800-631-17	Sequence 17, Appli
c	26	296	6.8	106323	10	US-09-803-661-3	Sequence 3, Appli
c	27	294.6	6.8	3083	10	US-09-764-869-1703	Sequence 1703, Ap
c	28	291.6	6.7	145831	10	US-09-969-708-79	Sequence 79, Appli
c	29	291.6	6.7	145831	10	US-09-954-456-2116	Sequence 2116, Ap
c	30	291.2	6.7	36159	12	US-10-135-687-3	Sequence 3, Appli
c	31	290.8	6.7	6892	10	US-09-764-869-1720	Sequence 1720, Ap
c	32	290	6.7	684973	10	US-09-263-959-1	Sequence 1, Appli
c	33	289	6.7	46050	10	US-09-820-003A-3	Sequence 3, Appli
c	34	288.4	6.6	40645	9	US-10-216-441-3	Sequence 3, Appli
c	35	288.4	6.6	40645	10	US-09-818-656A-3	Sequence 3, Appli
c	36	288.2	6.6	32183	10	US-09-764-869-1494	Sequence 1494, Ap
c	37	287.6	6.6	57130	10	US-09-835-081-3	Sequence 3, Appli
c	38	286.8	6.6	5251	10	US-09-764-869-1431	Sequence 1431, Ap
c	39	286.2	6.6	2550	10	US-09-880-107-2131	Sequence 2131, Ap
c	40	286.2	6.6	5095	10	US-09-824-637-3	Sequence 3, Appli
c	41	285	6.6	17509	10	US-09-880-107-2097	Sequence 2097, Ap
c	42	282.6	6.5	174424	10	US-09-967-768A-314	Sequence 314, App
c	43	282	6.5	27332	10	US-09-764-847-1555	Sequence 1555, Ap
c	44	281	6.5	32204	10	US-09-764-855-327	Sequence 327, App
c	45	280.8	6.5	34641	10	US-09-954-456-1110	Sequence 1110, Ap

ALIGNMENTS

RESULT 1
US-09-924-396B-17
; Sequence 17, Application US/09924396B
; Patent No. US20020165349A1
; GENERAL INFORMATION:
; APPLICANT: Kirsch, Wolff
; APPLICANT: Lennart, Anton
; APPLICANT: Kelln, Wayne
; APPLICANT: Kang, Dae-Kyung
; APPLICANT: Levine, Rodney
; APPLICANT: Rouault, Tracey
; TITLE OF INVENTION: IRON-REGULATING PROTEIN-2 (IRP-2) IS
; FILE REFERENCE: LOMAU.140A
; CURRENT APPLICATION NUMBER: US/09/924.396B
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 60/222,863
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 2867
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-09-924-396B-17

Query Match 22.8%; Score 990; DB 9; Length 2867;
Best Local Similarity 100.0%; Pred. No. 8.3e-169;
Matches 990; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 TACCCTATTGAACATTAAATGACAGTTTCACATAGAAGTTCTTCGATGTATCTAAACTT 93

Db 1 TACCCTATTGAACATTAAATGACAGTTTCACATAGAAGTTCTTCGATGTATCTAAACTT 60

QY 94 GGCACCAAGATGATGTTCTGCCTTACTCAATACGGGCTTTGTTGGAAGCTGCTGTACGA 153

Db 61 GGCACCAAGATGATGTTCTGCCTTACTCAATACGGGCTTTGTTGGAAGCTGCTGTACGA 120

QY 154 AATTGTGATGGCTTTTAAATGAAGAAGGAGATGTTATGAACATTTTAACTGGAACACC 213

Db 121 AATTGTGATGGCTTTTAAATGAAGAAGGAGATGTTATGAACATTTTAACTGGAACACC 180

214 AAACAAAGCAATGTTGAAGTCCCTTTTCCCTGCCGCTGTCTCTTCAAGATTTTACT 273
181 AAACAAAGCAATGTTGAAGTCCCTTTTCCCTGCCGCTGTCTCTTCAAGATTTTACT 240
274 GGAATACAGCAATGGTGGATTTTGTCTGTATGAGGAGGAGCAATCTTTGGAGGT 333
241 GGAATACAGCAATGGTGGATTTTGTCTGTATGAGGAGGAGCAATCTTTGGAGGT 300
334 GATCCTGAGAAAGTCCATCTGCTGTCCGACAGATCTTACAGTTGACCATCTTTTACAA 393
301 GATCCTGAGAAAGTCCATCTGCTGTCCGACAGATCTTACAGTTGACCATCTTTTACAA 360
394 ATTGACTTTCAGTAAATGTGCAATACAGAAATGCACCAATCTCGAGGTGTGACCTGCAG 453
361 ATTGACTTTCAGTAAATGTGCAATACAGAAATGCACCAATCTCGAGGTGTGACCTGCAG 420
454 AAAGCAGAAAGCTCTCTCCACTTAAAGTGCAGGCTTAGAAGCTTCCCTGCAGAGGCCAG 513
421 AAAGCAGAAAGCTCTCTCCACTTAAAGTGCAGGCTTAGAAGCTTCCCTGCAGAGGCCAG 480
514 ACTACCTGCCGAGGATCTGTGATCTGAGAACTAGGCCGAAACTCAGGAACATTTTCT 573
481 ACTACCTGCCGAGGATCTGTGATCTGAGAACTAGGCCGAAACTCAGGAACATTTTCT 540
574 TCGCAGATTGGAATACACCCATCCTGTCTCTTTTCAATTTGCAACCAAGTGCCTGAACCT 633
541 TCGCAGATTGGAATACACCCATCCTGTCTCTTTTCAATTTGCAACCAAGTGCCTGAACCT 600
634 GAACAGTGTAAAAAATCAGAAGTAGAATTCGCGAGAAATCGAGAGGCTTCAGTTT 693
601 GAACAGTGTAAAAAATCAGAAGTAGAATTCGCGAGAAATCGAGAGGCTTCAGTTT 660
694 TTTAAGTGGAGTCAAGAGTTTAAAGAAATGTGGCAGTATCCCTCGGAACTGGAAATG 753
661 TTTAAGTGGAGTCAAGAGTTTAAAGAAATGTGGCAGTATCCCTCGGAACTGGAAATG 720
754 GCTCATCAAAATAAATAGATTTAAGATTTGCAAGAGTGGTTTTGAGAAAAAGACCTCCTC 813
721 GCTCATCAAAATAAATAGATTTAAGATTTGCAAGAGTGGTTTTGAGAAAAAGACCTCCTC 780
814 TTCCAGACAGTGTAGTCGGACAGATTCACACATAAGCAGTGGTGAATGTTTAGGGATT 873
781 TTCCAGACAGTGTAGTCGGACAGATTCACACATAAGCAGTGGTGAATGTTTAGGGATT 840
874 CTGGGTGGGGGTGGAGGATTTGAAACAGAGCAGTTATGCTTGTGTCCTGCCAGTTTCT 933
841 CTGGGTGGGGGTGGAGGATTTGAAACAGAGCAGTTATGCTTGTGTCCTGCCAGTTTCT 900
934 CTACTTTACAGAGGTGGTGGATGCTGAGTTAAGTGAAGTGAATGTTTAGGGATT 993
901 CTACTTTACAGAGGTGGTGGATGCTGAGTTAAGTGAAGTGAATGTTTAGGGATT 960
994 TCCATAGATGTTGTTCTGGTATTACAAG 1023
961 TCCATAGATGTTGTTCTGGTATTACAAG 990

RESULT 2

US-09-924-396B-19
; Sequence 19, Application US/09924396B
; Patent No. US20020165349A1
; GENERAL INFORMATION:
; APPLICANT: Kirsch, Wolff
; APPLICANT: Lennart, Anton
; APPLICANT: Kellin, Wayne
; APPLICANT: Kang, Dae-Kyung
; APPLICANT: Rouault, Tracey
; APPLICANT: Levine, Rodney
; TITLE OF INVENTION: IRON-REGULATING PROTEIN-2 (IRP-2) IS
; FILE REFERENCE: LONAU.140A
; CURRENT APPLICATION NUMBER: US/09/924,396B
; CURRENT FILING DATE: 2002-04-25

; PRIOR APPLICATION NUMBER: 60/222,863
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 3770
; TYPE: DNA
; ORGANISM: Rattus No. US20020165349Alvegicus
US-09-924-396B-19

Query Match 20.2%; Score 877.2; DB 9; Length 3770;
Best Local Similarity 91.6%; Pred. No. 1.5e-148;
Matches 940; Conservative 0; Mismatches 83; Indels 3; Gaps 1;

QY 1 ATGAGCGCCCAAAAGCAGGATACGCTTTTGAGTACCTTATTGAAACATTAATGACAGT 60
DB 12 ATGAGCTCCCAAGTCAGGATACACCTTTGAGTACCTTATTGAAACATTAATGACAGT 71
QY 61 TCACATAAGAAGTTCCTCGATGATCTAACTT---GGCACCAAGTATGATGTTCTGCCT 117
DB 72 TCACAGAAGAAGTTCCTCAATGTACCTAACTTTGGAGGCCAACAGTATGATATCTGCCT 131
QY 118 TACTCAATACGGCTCTGTGGAGCTGCTGTACGAAATTTGTATGGCTTTTAAATGAAG 177
DB 132 TACTCAATACGGCTCTGTGGAGCTGCTGTACGAAATTTGTATGGATTTTAAATGAAG 191
QY 178 AAGGAAGATGTTATGAACATTTTAGACTTGAAACCAACCAAGCAATGTTGAAGTGCCC 237
DB 192 AAGGAAGATGTTATAATATTTTGGACTGGAACCAACCAAGCAATGTTGAAGTGCCC 251
QY 238 TTTTTCCTCCCGTGTCTTCTTCAAGATTTTACGGAATACCACCAATGTTGGATTTT 297
DB 252 TTTTTCCTCCCGTGTCTTCTTCAAGATTTTCACTGGAATACCGCAATGTTGGATTTT 311
QY 298 GCTGCTATGAGGAGGAGGAGTAACTCTTGGAGGTGATCCTCAGAAAGTCCATCCTGCT 357
DB 312 GCTGCTATGAGGAGGAGGAGTAACTCTTGGAGGTGATCCTCAGAAAGTCCATCCTGCT 371
QY 358 TGTCCGACAGATCTTACAGTTGACCAATCTTTTACAAATGACATTCAGTAAATGTCGAATA 417
DB 372 TGTCCAAACAGATCTCAGAGTTGACCACTCTTTACAGATTTGACTTCAATGTCGAATA 431
QY 418 CAGATCCACCAATCCTGGAGGTGGTACCTCGCAAGAGCAGAAAGCTCTCTCCACTT 477
DB 432 CAGAAATCCACCAATCCTGGAGGTGGTACCTCAGAAAGCAGAAAGCTCTCTCCACTT 491
QY 478 AAAGTGCAGCCTAAGAGCTTCCCTGCAGAGGCCAGACTACCTGCCGAGGATCTTGTGAT 537
DB 492 AAAGTGCAGCCTAAGAGCTTCCCTGCAGAGGCCAGACTACCTGCCGAGGATCTTGTGAT 551
QY 538 TCTGGAGAACTAGCCGAAACTCAGGAACATTTTCTTCCGAGATTTGAGAAATACCCCATC 597
DB 552 TCTGGAGAACTAAGCCGAAACTCAGGAACATTTTCTTCCGAGATTTGAGAAATACCCCATC 611
QY 598 CTGTGTCCTTTTCAATTTGCAACAGTCCCTGAACTGAAACAGTGTAAATAATCAAGAA 657
DB 612 CTGTGTCCTTTTCAATTTGCAACAGTCCCTGAACTGAAACAGTGTAAATAATCAAGAA 671
QY 658 GTAGAATTCGCGAGAAATCGAGAGAGGCTTTCAGTTTAAAGTGGAGTTTCAAGAGTTTAA 717
DB 672 GTAGAATTCGCGAGAAATCGAGAGAGGCTTTCAGTTTAAAGTGGAGTTCAGGAGCTTTT 731
QY 718 AAGAATCTGCGAGTATCCCTCTCGAACTGGAAATGAGTCTCATCAATAAATAGTATATAT 777
DB 732 AAGAATCTGCGAGTATCCCTCTCGAACTGGAAATGAGTCTCATCAAGTGAACCTTAGAACAT 791
QY 778 TTGTCAAGAGTGGTTTTTGAAGAAAGAGCTCTCTTCCACAGACAGTGTAGTCGCACA 837
DB 792 TTGTCAAGAGTGGTTTTTGAAGAAAGAGCTCTCTTCCACAGACAGCTGAATTTGGCACA 851
QY 838 GATTTCACACATACAGTGGTCAATGTTTAGGAGTCTCGGGGTGGGGGTGGAGGCAAT 897
DB 852 GATTTCATATACCATGGTGAATGGATTGGGAATCTTGGTGGGAGTGGAGGCAAT 911


```

; ORGANISM: human
US-09-933-267A-1

Query Match      7.3%; Score 317.4; DB 10; Length 465237;
Best Local Similarity 93.5%; Pred. No. 1.1e-47;
Matches 330; Conservative 1; Mismatches 22; Indels 0; Gaps 0;

QY 2387 TCTTTTCATCCCTAAANAATAAACCATAGGCGGCGCGGTGGCTACGCCCTGTAATCCCA 2446
DB 47735 TATTTCCTCTAAGGACAAAGAGCGCAGTGCGCGCGCGGTGGCTACGCCCTGTAATCCCA 47676
QY 2447 GCACCTTTGGGAGGCGGAGCGCGGTGGATCATGAGGTTCAGGAGATCGAGACCATCCTGGCT 2506
DB 47675 GCACCTTTGGGAGGCGGAGCGCGGTGGATCATGAGGTTCAGGAGATCGAGACCATCCTGGCT 47616
QY 2507 AACAGGTGAACCCCGTCTCTACTATAAATAACAAAAAATAGCCGGCGCGGTGGCGGG 2566
DB 47615 AACAGGTGAACCCCGTCTCTACTATAAATAACAAAAAATAGCCGGCGCGGTGGCGGG 47556
QY 2567 CGCCTGTAGTCCCGAGCTTCTCGGAGGCTGAGCAGGAGATGCGTGAACCCGGGAGGC 2626
DB 47555 CGCCTGTAGTCCCGAGCTTCTCGGAGGCTGAGCAGGAGATGCGTGAACCCGGGAGGC 47496
QY 2627 GGAGCTTGCAGTGAGCCCAAGATTGGCCACTGCAGTCCGCGAGTCCGCGCTGGCGCAGCA 2686
DB 47495 GGAGCTTGCAGTGAGCCGAGATTGGCCACTGCAGTCCGCGAGTCCGCGCTGGCGCAGCA 47436
QY 2687 GCGAGACTCTGTYTCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAGGA 2739
DB 47435 GCGAGACTCCGCTCTCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAGGA 47383

RESULT 5
US-09-949-654-3/c
; Sequence 3, Application US/09949654
; Patent No. US20020127644A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000817
; CURRENT APPLICATION NUMBER: US/09/949,654
; PRIOR FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: 60/231,572
; PRIOR FILING DATE: 2000-09-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 368004
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(368004)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-654-3

Query Match      7.2%; Score 311.8; DB 10; Length 368004;
Best Local Similarity 90.7%; Pred. No. 1e-46;
Matches 331; Conservative 1; Mismatches 33; Indels 0; Gaps 0;

QY 2370 AATATCTAATCTGTGTTCTTCTTCATCTAAANAATAAACCATAGGCGGCGCGGTGGC 2429
DB 286574 AGTGTTTGTCTTATTTTGTGGGAATAAATATGTACCAATAGCCGGCGCGGTGGC 286515
QY 2430 TCACGCGCTTAATCCAGACATTTTGGGAGCGGAGCGCGGTGGATCATGAGTTCAGGAGA 2489
DB 286514 TCACGCGCTTAATCCAGACATTTTGGGAGCGGAGCGCGGTGGATCATGAGTTCAGGAGA 286455
QY 2490 TCGAGACCATCTGGGTACACAGGTGAACCCCGTCTCTACTAAAAATACAAAAAATTAG 2549
DB 286454 TCGAGACCATCTGGGTACACAGGTGAACCCCGTCTCTACTAAAAATACAAAAAATTAG 286395
```

```

QY 2550 CCGGGCGCGGTGGCGGCGCCTTGTAGTCCAGCTTCTCGGGAGGTGAGGAGGAATG 2609
DB 286394 CCGGGCGCGGTGGCGGCGCCTTGTAGTCCAGCTTCTCGGGAGGTGAGGAGGAATG 286335
QY 2610 GCGTGAACCCGGGAGGCGGAGCTTGCAGTGAGCCCAAGATTGGCCCACTGCAGTCCGCACT 2669
DB 286334 GCGTGAACCCGGGAGGCGGAGCTTGCAGTGAGCCCAAGATTGGCCCACTGCAGTCCGCACT 286275
QY 2670 CCGGCTGTGGCGACAGAGAGAGACTCTGTYTCAAAAAAATAAAAAAATAAAAAAATAAAAAA 2729
DB 286274 CCGGCTGTGGCGACAGAGAGAGACTCTGTYTCAAAAAAATAAAAAAATAAAAAAATAAAAAA 286215
QY 2730 TAAAT 2734
DB 286214 ATAAT 286210

RESULT 6
US-09-764-870-598/c
; Sequence 598, Application US/09764870
; Patent No. US20020042386A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT214
; CURRENT APPLICATION NUMBER: US/09/764,870
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 598
; LENGTH: 20210
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-870-598

Query Match      7.2%; Score 311; DB 10; Length 20210;
Best Local Similarity 88.9%; Pred. No. 7.5e-47;
Matches 335; Conservative 1; Mismatches 41; Indels 0; Gaps 0;

QY 2392 CATCCTAAAAATAAACCATAGGCGGCGCGGTGGCTACGCCCTGTATCCAGCACT 2451
DB 16817 CATGCATGAAAGTCCCGGCGGCGGTGGCTACGCCCTGTATCCAGCACT 16758
QY 2452 TTGGGAGCGGAGCGGCTGGATCATGAGTCAGGAGATCGAGACCATCTTGGCTAACAA 2511
DB 16757 TTGGGAGCGGAGCGGCTGGATCATGAGTCAGGAGATCGAGACCATCTTGGCTAACAA 16698
QY 2512 GGTGAAACCCCGTCTCTACTATAAATAACAAAAATAGCGGCGCGGTGGCGGCGCCT 2571
DB 16697 GGTGAAACCCCGTCTCTACTATAAATAACAAAAATAGCGGCGCGGTGGCGGCGCCT 16638
QY 2572 GTAGTCCAGCTTCTCGGAGGCTGAGGAGAGATGGGTGAACCCGGGAGCGGAGC 2631
DB 16637 GTAGTCCAGCTTCTCGGAGGCTGAGGAGAGATGGGTGAACCCGGGAGCGGAGC 16578
QY 2632 TTGCAGTGAGCAAGATTGGCCCACTGCAGTCCGCGAGTCCGCGCTGGGCGACAGCGAG 2691
DB 16577 TTGCAGTGAGCGGAGATTGGCCCACTGCAGTCCGCGAGTCCGCGCTGGGCGACAGCGAG 16518
QY 2692 ACTCTGTYTCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATCTTTA 2751
DB 16517 ACTCCGTCTCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATATG 16458
QY 2752 CACTTAGGTTTGAGTT 2768
DB 16457 CATGTTACCTTTGAGCT 16441
```

```

RESULT 7
US-10-095-407-17
; Sequence 17, Application US/10095407
```

[illegible]

|||||
Db 15027 AGTCCAGCTACTCGGAGGTGAGCAGAGATGGGTGAACCCGGAGCGAGCTT 14968
Qy 2634 CGAGTGAGCAGAGATTGCCCACTGCAGTCCGCGCTCGGCGTGGGACAGAGCGAGAC 2693
|||||
Db 14967 CGAGTGAGCGGAGATTGCCCACTGCAGTCCGCGCTCGGCGTGGGACAGAGCGAGAC 14908
Qy 2694 TCTGTYTCAAAAAAAAAAAAAAAAAAAAAA 2726
|||||
Db 14907 TCCGCTCTCAAAAAAAAAAAAAAAAAAAAAA 14875

RESULT 9

US-09-764-868-1282
; Sequence 1282, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1282
; LENGTH: 13873
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (10983)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-868-1282

Query Match 7.1%; Score 309.6; DB 9; Length 13873;
Best Local Similarity 91.3%; Pred. No. 1.2e-46;
Matches 327; Conservative 1; Mismatches 30; Indels 0; Gaps 0;

Qy 2383 TGTTTCTTTCATCTTAAATAAATTAACCATAGCCGGCGCGGTGCTCACCGCTGTAAT 2442
Db 2156 TGGATTCTTCCCTTAAAGAAAGCACTGTGGCGCGGCGAGTGGCTCACGCGTGAAT 2215
Qy 2443 CCAGCAGCTTTGGAGCCGAGCGGTGATCATGAGTTCAGGAGATCGAGACCATCCT 2502
Db 2216 CCAGCAGCTTTGGAGCCGAGCGGTGATCATGAGTTCAGGAGATCGAGACCATCCT 2275
Qy 2503 GGCTAACAGGTGAACCCCGTCTCTACTAAAAATACAAAAATTAAGCGCGCGGTGG 2562
Db 2276 GGCTAACAGGTGAACCCCGTCTCTACTAAAAATACAAAAATTAAGCGCGCGGTGG 2335
Qy 2563 CGGCGCGCTGTAGTCCAGCTTCTCGGAGGCTGAGCAGGAGAATGGGTGAACCCGGG 2622
Db 2336 CGGCGCGCTGTAGTCCAGCTTCTCGGAGGCTGAGCAGGAGAATGGGTGAACCCGGG 2395
Qy 2623 AGGCGGAGCTTGAGTGAGCAGCAAGATTGCCCACTGCAGTCCGCGCTGGGCGA 2682
Db 2396 AAGCGGAGCTTGAGTGAGCAGCAAGATTGCCCACTGCAGTCCGCGCTGGGCGA 2455
Qy 2683 CAGAGCGAGCTCTGTTCAAAAAATAAAAAAAAAAAAAAAAAAACCAATATGAGNA 2740
Db 2456 CAGAGCGAGCTCTGTTCAAAAAATAAAAAAAAAAAAAAAAAAACCAATATGAGNA 2513

RESULT 10

US-09-764-868-1424/c
; Sequence 1424, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868

; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1424
; LENGTH: 27007
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-868-1424

Query Match 7.1%; Score 309.2; DB 9; Length 27007;
Best Local Similarity 96.9%; Pred. No. 1.7e-46;
Matches 314; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

Qy 2413 AGCGCGGCGCGGTGCTCACGCTGTAAATCCAGCACTTTGGAGCGCGAGCGGTGG 2472
Db 11724 AGCGCGGCGCGGTGCTCACGCTGTAAATCCAGCACTTTGGAGCGCGAGCGGTGG 11665
Qy 2473 ATCATGAGGTGAGGATCGAGACCATCCTGGCTAAACAAGTGAACCCGCTCTACTA 2532
Db 11664 ATCATGAGGTGAGGATCGAGACCATCCTGGCTAAACAAGTGAACCCGCTCTACTA 11605
Qy 2533 AAAATACAAAAAATTAGCGCGCGGTGCGCGCGCTGTAGTCCAGCTTCTCGGGAG 2592
Db 11604 AAAATACAAAAAATTAGCGCGCGGTGCGCGCGCTGTAGTCCAGCTTCTCGGGAG 11545
Qy 2593 GCTGAGCAGGAGAAATGCGTGAACCCGCGGAGCGGAGTTGAGTGAGCCAAATTCG 2652
Db 11544 GCTGAGCAGGAGAAATGCGTGAACCCGCGGAGCGGAGTTGAGTGAGCCAAATTCG 11485
Qy 2653 CCATGAGTCCGAGTCCGCGCTGGCGCGAGCAGAGCGAGCTGTGTCAAAAAA 2712
Db 11484 CCATGAGTCCGAGTCCGCGCTGGCGCGAGCAGAGCGAGCTGTGTGTCAAAAAA 11425
Qy 2713 AAAAAAATAAACCATTAATGA 2736
Db 11424 AAAAAAATAAACCATTAATGA 11401

RESULT 11

US-10-095-407-16/c
; Sequence 16, Application US/10095407
; Patent No. US20020164330A1
; GENERAL INFORMATION:
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
; FILE REFERENCE: 09404/052001
; CURRENT APPLICATION NUMBER: US/10/095,407
; CURRENT FILING DATE: 2002-03-11
; PRIOR APPLICATION NUMBER: US 60/091,650
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: US 60/054,646
; PRIOR FILING DATE: 1997-08-04
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 152331
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)....(152331)
; OTHER INFORMATION: n = A,T,C or G
US-10-095-407-16

Query Match 7.1%; Score 309; DB 9; Length 152331;
Best Local Similarity 96.3%; Pred. No. 2.7e-46;
Matches 315; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

Qy 2397 TAAAAATAAACCATATAGCCGGCGCGGTGCTCAGCCTGTAAATCCAGCAGCTTGGG 2456
Db 114670 TTAATAATGCCACTCTTGGCGCGCGGTGCTCAGCCTGTAAATCCAGCAGCTTGGG 114611

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 28, 2003, 05:39:35 ; Search time 8804 seconds
(without alignments)
12402.751 Million cell updates/sec

Title: US-09-730-559B-7
Perfect score: 4343
Sequence: 1 atggacgcccccaagagg.....attcacagcagtcctctct 4343

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 24791104 seqs, 12571243825 residues 49582208
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_Main:*

1: /cgn2_6/ptodata/1/pna/PCTUS_COMB.seq:*

2: /cgn2_6/ptodata/1/pna/US06_COMB.seq:*

3: /cgn2_6/ptodata/1/pna/US07_COMB.seq:*

4: /cgn2_6/ptodata/1/pna/US08_COMB.seq:*

5: /cgn2_6/ptodata/1/pna/US081_COMB.seq:*

6: /cgn2_6/ptodata/1/pna/US082_COMB.seq:*

7: /cgn2_6/ptodata/1/pna/US083_COMB.seq:*

8: /cgn2_6/ptodata/1/pna/US084_COMB.seq:*

9: /cgn2_6/ptodata/1/pna/US085_COMB.seq:*

10: /cgn2_6/ptodata/1/pna/US086_COMB.seq:*

11: /cgn2_6/ptodata/1/pna/US087_COMB.seq:*

12: /cgn2_6/ptodata/1/pna/US088_COMB.seq:*

13: /cgn2_6/ptodata/1/pna/US089_COMB.seq:*

14: /cgn2_6/ptodata/1/pna/US090_COMB.seq:*

15: /cgn2_6/ptodata/1/pna/US091_COMB.seq:*

16: /cgn2_6/ptodata/1/pna/US092_COMB.seq:*

17: /cgn2_6/ptodata/1/pna/US093_COMB.seq:*

18: /cgn2_6/ptodata/1/pna/US094_COMB.seq:*

19: /cgn2_6/ptodata/1/pna/US095A_COMB.seq:*

20: /cgn2_6/ptodata/1/pna/US095B_COMB.seq:*

21: /cgn2_6/ptodata/1/pna/US095C_COMB.seq:*

22: /cgn2_6/ptodata/1/pna/US095D_COMB.seq:*

23: /cgn2_6/ptodata/1/pna/US096A_COMB.seq:*

24: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:*

25: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:*

26: /cgn2_6/ptodata/1/pna/US096D_COMB.seq:*

27: /cgn2_6/ptodata/1/pna/US096E_COMB.seq:*

28: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:*

29: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:*

30: /cgn2_6/ptodata/1/pna/US097C_COMB.seq:*

31: /cgn2_6/ptodata/1/pna/US098A_COMB.seq:*

32: /cgn2_6/ptodata/1/pna/US098B_COMB.seq:*

33: /cgn2_6/ptodata/1/pna/US098C_COMB.seq:*

34: /cgn2_6/ptodata/1/pna/US099A_COMB.seq:*

35: /cgn2_6/ptodata/1/pna/US099B_COMB.seq:*

36: /cgn2_6/ptodata/1/pna/US099C_COMB.seq:*

37: /cgn2_6/ptodata/1/pna/US099D_COMB.seq:*

38: /cgn2_6/ptodata/1/pna/US100A_COMB.seq:*

39: /cgn2_6/ptodata/1/pna/US100B_COMB.seq:*

40: /cgn2_6/ptodata/1/pna/US101A_COMB.seq:*

41: /cgn2_6/ptodata/1/pna/US101B_COMB.seq:*

42: /cgn2_6/ptodata/1/pna/US102A_COMB.seq:*

43: /cgn2_6/ptodata/1/pna/US102B_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	4341.8	100.0	4343	29	US-09-730-559B-7
2	3248.4	74.8	62724	68	US-60-243-468-591
3	3248.4	74.8	62724	68	US-60-243-742-99
4	1023	23.6	3280	25	US-09-654-935A-257
5	1023	23.6	3290	1	PCT-US01-08656-4031
6	1023	23.6	4684	17	US-09-359-922-8303
7	1023	23.6	4684	17	US-09-359-922-8303
8	1023	23.6	4684	34	US-09-919-002-8303
9	1019.8	23.5	4194	18	US-09-471-275-1719
10	1019.8	23.5	4194	18	US-09-496-914A-7145
11	1019.8	23.5	4194	22	US-09-560-875A-7145
12	1019.8	23.5	4729	76	US-60-324-185-26664
13	1019.8	23.5	4871	77	US-60-339-453-290
14	1004	23.1	5034	77	US-60-339-453-63
15	990	22.8	2867	34	US-09-924-396B-17
16	990	22.8	2867	66	US-60-222-863-17
17	990	22.8	3558	16	US-09-205-070-12807
18	990	22.8	3558	17	US-09-340-623-12807
19	990	22.8	3558	33	US-09-898-888-12807
20	990	22.8	3558	33	US-09-898-888A-12807
21	877.2	20.2	3770	34	US-09-924-396B-19
					Sequence 7, Appl
					Sequence 591, Appl
					Sequence 99, Appl
					Sequence 257, Appl
					Sequence 4031, Ap
					Sequence 8303, Ap
					Sequence 8303, Ap
					Sequence 8303, Ap
					Sequence 1719, Ap
					Sequence 7145, Ap
					Sequence 26664, A
					Sequence 290, Appl
					Sequence 63, Appl
					Sequence 17, Appl
					Sequence 12807, A
					Sequence 12807, A
					Sequence 12807, A
					Sequence 12807, A

22	877.2	20.2	3770	66	US-60-222-863-19	Sequence 19, Appl
23	767.8	17.7	814	29	US-09-736-119-1624	Sequence 1624, Ap
24	654.8	15.1	5601	80	US-60-360-207-5266	Sequence 5266, Ap
25	526.2	12.1	605	30	US-09-760-475-125	Sequence 125, App
26	526.2	12.1	605	30	US-09-760-475-1331	Sequence 1331, Ap
27	506.6	11.7	561	24	US-09-637-889-4744	Sequence 4744, Ap
28	304	11.6	597	29	US-09-736-119-350	Sequence 350, App
c	29	494.4	581	1	PCT-US02-25766-12391	Sequence 12391, A
30	469.6	10.8	608	30	US-09-785-276A-56174	Sequence 56174, A
31	435.8	10.0	508	65	US-60-213-362-6374	Sequence 6374, A
32	435.8	10.0	508	76	US-60-324-185-27788	Sequence 27788, A
33	381	8.8	391	32	US-09-834-366-37869	Sequence 37869, A
34	381	8.8	391	63	US-60-197-873-37869	Sequence 37869, A
35	381	8.8	414	19	US-09-528-409-111255	Sequence 111255,
36	381	8.8	414	35	US-09-933-524-111255	Sequence 111255,
37	381	8.8	414	35	US-09-933-524A-111255	Sequence 111255,
38	377.2	8.7	440	18	US-09-489-036-34205	Sequence 34205, A
39	377.2	8.7	440	35	US-09-943-143-34205	Sequence 34205, A
40	372	8.6	409	19	US-09-528-409-49242	Sequence 49242, A
41	372	8.6	409	35	US-09-933-524-49242	Sequence 49242, A
42	372	8.6	409	35	US-09-933-524A-49242	Sequence 49242, A
43	369	8.5	407	19	US-09-528-409-49241	Sequence 49241, A
44	369	8.5	407	35	US-09-933-524-49241	Sequence 49241, A
45	369	8.5	407	35	US-09-933-524A-49241	Sequence 49241, A

ALIGNMENTS

RESULT 1
US-09-730-559B-7
; Sequence 7, Application US/09730559B
; GENERAL INFORMATION:
; APPLICANT: ISHIWATA, TETSUYOSHI
; APPLICANT: SAKURADA, MIKIO
; APPLICANT: KANABATA, AYAKO
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: NISHI, TATSUNARI
; APPLICANT: KUGA, TETSURO
; APPLICANT: SAWADA, SHIGEMASA
; APPLICANT: TAKEI, MASAMI
; APPLICANT: SHIBATA, KENJI
; APPLICANT: FURUYA, AKIKO
; TITLE OF INVENTION: Iga NEPHROPATHY-ASSOCIATED GENE
; FILE REFERENCE: 766.21 CIP
; CURRENT APPLICATION NUMBER: US/09/730,559B
; CURRENT FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 4343
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1029)
US-09-730-559B-7

Query Match	100.0%	Score 4341.8;	DB 29;	Length 4343;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 4343;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	ATGGACGCCCAAGCAGGATACGCCCTTTGAGTACCTTATTTGAACATTAAATGACAGT	60	
Db	1	ATGGACGCCCAAGCAGGATACGCCCTTTGAGTACCTTATTTGAACATTAAATGACAGT	60	
Qy	61	TCACATAGAAGTCTTCGATGTATCTAAACTTGGCACCAGATGATGTCTGCCCTTAC	120	
Db	61	TCACATAGAAGTCTTCGATGTATCTAAACTTGGCACCAGATGATGTCTGCCCTTAC	120	
Qy	121	TCAATACGGGCTCTGTTGGAAGCTGCTGTACGAAATTTGTATGCTTTTTTAATGAAGAAG	180	
Db	121	TCAATACGGGCTCTGTTGGAAGCTGCTGTACGAAATTTGTATGCTTTTTTAATGAAGAAG	180	

Qy	181	GAAGATGTTATGAACATTTTGTAGCTGGAACCAAAACCAAGCAATGTTGAAGTGCCTTT	240	
Db	181	GAAGATGTTATGAACATTTTGTAGCTGGAACCAAAACCAAGCAATGTTGAAGTGCCTTT	240	
Qy	241	TTCCCTCCCGTGTCTTCTTCAAGATTTTACTGGAAATACCAGCAATGGTGGATTTTGTCT	300	
Db	241	TTCCCTCCCGTGTCTTCTTCAAGATTTTACTGGAAATACCAGCAATGGTGGATTTTGTCT	300	
Qy	301	GCATAGAGGAGCAGTGAAACCTCTGGAGGTGATCCCTCAGAAAGTCCATCCTGCTTCT	360	
Db	301	GCATAGAGGAGCAGTGAAACCTCTGGAGGTGATCCCTCAGAAAGTCCATCCTGCTTCT	360	
Qy	361	CCGACAGATCTTACAGTTGACCATTTCTTACAAATTTGACTTTCAGTAAATGTTGCAATACAG	420	
Db	361	CCGACAGATCTTACAGTTGACCATTTCTTACAAATTTGACTTTCAGTAAATGTTGCAATACAG	420	
Qy	421	ANTGCACCAATCCTGGAGTGTGACCTGCAGAAAGCAGGAAGCTCTCTCCACTTAA	480	
Db	421	ANTGCACCAATCCTGGAGTGTGACCTGCAGAAAGCAGGAAGCTCTCTCCACTTAA	480	
Qy	481	GTGAGGCTTAAGAGCTTCCCTGCAGAGCCAGACTACCTGCGGAGGATCTTGTGATTTCT	540	
Db	481	GTGAGGCTTAAGAGCTTCCCTGCAGAGCCAGACTACCTGCGGAGGATCTTGTGATTTCT	540	
Qy	541	GGAGAACTAGCGCGAAACTCAGGAACATTTTCTTCGAGATTGAGAAATACCCATCCTG	600	
Db	541	GGAGAACTAGCGCGAAACTCAGGAACATTTTCTTCGAGATTGAGAAATACCCATCCTG	600	
Qy	601	TGTCCTTTTCATTTGCAACAGTGCCTGAACCTGGAACAGTGTAAAAAATCAAGAAGTA	660	
Db	601	TGTCCTTTTCATTTGCAACAGTGCCTGAACCTGGAACAGTGTAAAAAATCAAGAAGTA	660	
Qy	661	GAATTCGGCAGAAATCGAGAGAGCTTCAGTCTTTTAAAGTGGAGTTCAGAGTCTTAAAG	720	
Db	661	GAATTCGGCAGAAATCGAGAGAGCTTCAGTCTTTTAAAGTGGAGTTCAGAGTCTTAAAG	720	
Qy	721	AATGTGGCAGTGTATCCTCCTGGAACCTGGAATGGCTCATCAATAAATTTAGAAATTTG	780	
Db	721	AATGTGGCAGTGTATCCTCCTGGAACCTGGAATGGCTCATCAATAAATTTAGAAATTTG	780	
Qy	781	TCAAGAGTGGTTTTTGAAGAAAGACCTCCCTCTCCAGACAGTGTAGTGGCAGAGAT	840	
Db	781	TCAAGAGTGGTTTTTGAAGAAAGACCTCCCTCTCCAGACAGTGTAGTGGCAGAGAT	840	
Qy	841	TCACATACAGATGTTGATGTTAGGATTCGCGGTGGGGGTTCGAGGCAATTTGAA	900	
Db	841	TCACATACAGATGTTGATGTTAGGATTCGCGGTGGGGGTTCGAGGCAATTTGAA	900	
Qy	901	ACAGAAGCAGTTATGCTTGGTCTGCCAGTTTCTCTTACTTTACCAGAGTGGTGGATGT	960	
Db	901	ACAGAAGCAGTTATGCTTGGTCTGCCAGTTTCTCTTACTTTACCAGAGTGGTGGATGT	960	
Qy	961	GAGTTAAGTGGGTCATCAAAACCTTTTGTACATCCATAGATGTTGTTGGTATTACA	1020	
Db	961	GAGTTAAGTGGGTCATCAAAACCTTTTGTACATCCATAGATGTTGTTGGTATTACA	1020	
Qy	1021	AAGCTAAGTTAAAGTTGTTGTTAGTCTTACTGAAACATTTATTTTATAAAAATTTG	1080	
Db	1021	AAGCTAAGTTAAAGTTGTTGTTAGTCTTACTGAAACATTTATTTTATAAAAATTTG	1080	
Qy	1081	AAGAGCTCTATGAGCAGGAGTTTGGGTTCAATTTACTGCATCCTCAGGCTCTTGACGTT	1140	
Db	1081	AAGAGCTCTATGAGCAGGAGTTTGGGTTCAATTTACTGCATCCTCAGGCTCTTGACGTT	1140	
Qy	1141	AGCCACATCATAGTTATCATAGTAATAACAACAACAGAGCATTTAGTTTGTACTAA	1200	
Db	1141	AGCCACATCATAGTTATCATAGTAATAACAACAACAGAGCATTTAGTTTGTACTAA	1200	
Qy	1201	TAAATACAAGAAATTTGTTGTTTACCTTATGTTAGCTCATTTAGTCTCTTATAACAAG	1260	
Db	1201	TAAATACAAGAAATTTGTTGTTTACCTTATGTTAGCTCATTTAGTCTCTTATAACAAG	1260	

QY 1261 CTGTGAGATGGATCTATTACTATTCTCATTTGTAACCTGTGAGAAACTAAGGTACAGTAG 1320
 DB 1261 CTGTGAGATGGATCTATTACTATTCTCATTTGTAACCTGTGAGAAACTAAGGTACAGTAG 1320
 QY 1321 GGTTTAGTGACTTACCAAGGGTGAAGGGCTGAGTATAAGGGGTAGAGCAAAAGATTCCA 1380
 DB 1321 GGTTTAGTGACTTACCAAGGGTGAAGGGCTGAGTATAAGGGGTAGAGCAAAAGATTCCA 1380
 QY 1381 GCGAGTCAGATCTTGAGTCATGCTTAACCAATATGCTTATTAAGTCCCTGTGTGCTTGA 1440
 DB 1381 GCGAGTCAGATCTTGAGTCATGCTTAACCAATATGCTTATTAAGTCCCTGTGTGCTTGA 1440
 QY 1441 ATAAACACTTCTGCTGACTACATATTTTCTCTCTTTTAACTTGAATTAATAAAAAAAT 1500
 DB 1441 ATAAACACTTCTGCTGACTACATATTTTCTCTCTTTTAACTTGAATTAATAAAAAAAT 1500
 QY 1501 GTTTAGCAAAAGTTGAWTGTGCTCTTTTAAATTAATTTAACTTGAATTAATAAAAAAAT 1560
 DB 1501 GTTTAGCAAAAGTTGAWTGTGCTCTTTTAAATTAATTTAACTTGAATTAATAAAAAAAT 1560
 QY 1561 CTCTACTAAGTAATGCTTTCAAAACATGGACTCTAGAAATGTCATATATCATTTTCTG 1620
 DB 1561 CTCTACTAAGTAATGCTTTCAAAACATGGACTCTAGAAATGTCATATATCATTTTCTG 1620
 QY 1621 TTGCGGTTTTAACATTTCTCTGGAATTAATTAATAAAATCTTCTCTGAAATTTTAAAA 1680
 DB 1621 TTGCGGTTTTAACATTTCTCTGGAATTAATTAATAAAATCTTCTCTGAAATTTTAAAA 1680
 QY 1681 TACTGGCTTCAGAACTTCAATACACACTGAGCTTGTAAAGCATATTAATACACAGGCT 1740
 DB 1681 TACTGGCTTCAGAACTTCAATACACACTGAGCTTGTAAAGCATATTAATACACAGGCT 1740
 QY 1741 CAGGATTTCTAGTGAACAAATATTTGTAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1800
 DB 1741 CAGGATTTCTAGTGAACAAATATTTGTAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1800
 QY 1801 AACTTTATTTAATGATTAATCTATTTTGTAAATGAATGTACCTCGAAATTTGTCCA 1860
 DB 1801 AACTTTATTTAATGATTAATCTATTTTGTAAATGAATGTACCTCGAAATTTGTCCA 1860
 QY 1861 CATATAATCCAAATTTGAGTCCCAATCTCAGCAATTTTGGTTAGATTAATGGTACGAAG 1920
 DB 1861 CATATAATCCAAATTTGAGTCCCAATCTCAGCAATTTTGGTTAGATTAATGGTACGAAG 1920
 QY 1921 CTTTCTGGATCTCCAGTGAAGAAATGATAATGCTCCCTCTCAGCAATTTGGTATTGA 1980
 DB 1921 CTTTCTGGATCTCCAGTGAAGAAATGATAATGCTCCCTCTCAGCAATTTGGTATTGA 1980
 QY 1981 TCCTTCTTCCCTAATAGAAAAGAAATTTGGCACTTAGAGAAATTAATGATTCACCGAT 2040
 DB 1981 TCCTTCTTCCCTAATAGAAAAGAAATTTGGCACTTAGAGAAATTAATGATTCACCGAT 2040
 QY 2041 GATACCAAAAGATCAAGTAGTAATTTGGGAATTCAGAGATTAATTCAGAGAAAGAG 2100
 DB 2041 GATACCAAAAGATCAAGTAGTAATTTGGGAATTCAGAGATTAATTCAGAGAAAGAG 2100
 QY 2101 TATCCATTAATGTTTTACAGAAATCAATCTTTTACATTTAGACATCTGAAACCTAACGC 2160
 DB 2101 TATCCATTAATGTTTTACAGAAATCAATCTTTTACATTTAGACATCTGAAACCTAACGC 2160
 QY 2161 TGCTTTTATAGCTTCTCTAGCTGTTTTTCTGACAATATTAATGATTTTGTGTTTTGACAT 2220
 DB 2161 TGCTTTTATAGCTTCTCTAGCTGTTTTTCTGACAATATTAATGATTTTGTGTTTTGACAT 2220
 QY 2221 TTTAGTTTAAATGTTTTAAAAAATTAATCTATATATATATGTTTACATTTAATGAATATGA 2280
 DB 2221 TTTAGTTTAAATGTTTTAAAAAATTAATCTATATATATGTTTACATTTAATGAATATGA 2280
 QY 2281 TTACTTCTTTTTGAGATCCCTGTTCCCAATTTGTGATCCTTATAGGAATAATCCCTGATTGT 2340
 DB 2281 TTACTTCTTTTTGAGATCCCTGTTCCCAATTTGTGATCCTTATAGGAATAATCCCTGATTGT 2340
 QY 2341 TTTTGTGATGAGCAGCATTTGGTTGTAAATATCTAATCTGTTTCTTTTCATCCCTAAA 2400

DB 2341 TTTTGTGATGAGCAGCATTTGGTTGTAAATATCTAATCTGTGTTCTTTTCATCCTAAA 2400
 QY 2401 AAATAAAACCATAGCGGGCGGGTGGCTCAAGCTGTAAATCCCAAGCACTTTGGGAGGC 2460
 DB 2401 AAATAAAACCATAGCGGGCGGGTGGCTCAAGCTGTAAATCCCAAGCACTTTGGGAGGC 2460
 QY 2461 CGAGGCGGGTGGATCATGAGGTGAGGATCGAGACCATCTGGCTTAAACAAGGTGAAC 2520
 DB 2461 CGAGGCGGGTGGATCATGAGGTGAGGATCGAGACCATCTGGCTTAAACAAGGTGAAC 2520
 QY 2521 CCCTCTCTACTAAAAATACAAAAATTAGCGGGCGGGTGGCGGGCTCTAGTCCCA 2580
 DB 2521 CCCTCTCTACTAAAAATACAAAAATTAGCGGGCGGGTGGCGGGCTCTAGTCCCA 2580
 QY 2581 GCTTCTGGGAGGCTGAGGAGAGAAATGGGTGAACCCGGGAGGGAGCTTGCAGTGA 2640
 DB 2581 GCTTCTGGGAGGCTGAGGAGAGAAATGGGTGAACCCGGGAGGGAGCTTGCAGTGA 2640
 QY 2641 GCCAAGATTCGCGCACTGCGAGTCCGCGCTCCGCGCTGGCGGACAGAGCAGACTCTGTT 2700
 DB 2641 GCCAAGATTCGCGCACTGCGAGTCCGCGCTCCGCGCTGGCGGACAGAGCAGACTCTGTT 2700
 QY 2701 CAAAAAATAAAAAAATAAAAAACCATAAATAGAGAAACGCATCTTTACACTTAGG 2760
 DB 2701 CAAAAAATAAAAAAATAAAAAACCATAAATAGAGAAACGCATCTTTACACTTAGG 2760
 QY 2761 TTTGAGTCTCTGATCTATAAAAAAGGTTTGGATTAAGTATCCCTGGCACTTATAAAA 2820
 DB 2761 TTTGAGTCTCTGATCTATAAAAAAGGTTTGGATTAAGTATCCCTGGCACTTATAAAA 2820
 QY 2821 TGTTAGGCTTAAATATTTATAGATCAGGATAGTTTCAATCTTAGTCGCTCCTTAG 2880
 DB 2821 TGTTAGGCTTAAATATTTATAGATCAGGATAGTTTCAATCTTAGTCGCTCCTTAG 2880
 QY 2881 TCCTCTTCTCTATACCAATCTGAGACCAATTTACAAATTTAGAAAACAAATAAAGTTGAAAC 3000
 DB 2881 TCCTCTTCTCTATACCAATCTGAGACCAATTTACAAATTTAGAAAACAAATAAAGTTGAAAC 3000
 QY 3001 CACATGTTAAACAAATTTACCAAGTGGATTTGGCTGTGATTTAAAGATGCTGTAACAT 3060
 DB 3001 CACATGTTAAACAAATTTACCAAGTGGATTTGGCTGTGATTTAAAGATGCTGTAACAT 3060
 QY 3061 TTGGCCAGTAGTTAATTTGAAAAATGTTTATAGCCCAATATATAATTTTATTTAAA 3120
 DB 3061 TTGGCCAGTAGTTAATTTGAAAAATGTTTATAGCCCAATATATAATTTTATTTAAA 3120
 QY 3121 TATACAGTTTCATCAGTCTATTAGTATTTCATTAAGTCTAAGATGCCATCAGTGGTAGC 3180
 DB 3121 TATACAGTTTCATCAGTCTATTAGTATTTCATTAAGTCTAAGATGCCATCAGTGGTAGC 3180
 QY 3181 AAACACCACTGTTTATGCACTCTAAGAAAGAAATAAAGGCTGTGTCAGTGGCTCACA 3240
 DB 3181 AAACACCACTGTTTATGCACTCTAAGAAAGAAATAAAGGCTGTGTCAGTGGCTCACA 3240
 QY 3241 CCTGTGGGAGCCCAAGCAGGAGCATCACTTTGAGGCCAGAAAGTTCAAGACCAACCTGGTC 3300
 DB 3241 CCTGTGGGAGCCCAAGCAGGAGCATCACTTTGAGGCCAGAAAGTTCAAGACCAACCTGGTC 3300
 QY 3301 ACATTTGTAGACCTCTCTACAAAAAATAAAGTTTAAAAATAGTGGTGGGCGGTG 3360
 DB 3301 AACATTTGTAGACCTCTCTACAAAAAATAAAGTTTAAAAATAGTGGTGGGCGGTG 3360
 QY 3361 GCACATGCTGTATGTTCCAGCTACTCTGAGGCTAAGGTGGGAGGATTCCTAGACCCAG 3420
 DB 3361 GCACATGCTGTATGTTCCAGCTACTCTGAGGCTAAGGTGGGAGGATTCCTAGACCCAG 3420
 QY 3421 GTTTTGAAGCTGCAATGAGCTGTGACCAACCACTCTGCGCTCCAGCGTGGCAACAGAGT 3480

```
Db 3421 GTGTTGGAAGCTGCAATGAGCTGTGACCAACACCACTCGGCTCCAGCGTGGGCAACAGAT 3480
Qy 3481 GAGACCCCTGTTCTAAAGAAAGAAAGAAAGGGGTGCCACCTAAACAGACACACTAT 3540
Db 3481 GAGACCCCTGTTCTAAAGAAAGAAAGAAAGGGGTGCCACCTAAACAGACACACTAT 3540
Qy 3541 TGAGTTGAGGTACCCCTGATTTCAAGACATGAAATGTTAAATATAGCCACCTTGAAGCT 3600
Db 3541 TGAGTTGAGGTACCCCTGATTTCAAGACATGAAATGTTAAATATAGCCACCTTGAAGCT 3600
Qy 3601 TTCAGSCCCCTTCTACCCCTGAATTAACAGTGACATTTGGACGAGCTCTCTTTACTTCT 3660
Db 3601 TTCAGSCCCCTTCTACCCCTGAATTAACAGTGACATTTGGACGAGCTCTCTTTACTTCT 3660
Qy 3661 TATCTTAAATACCCCAACCAAGAGATGAGTTCATTCAATAGGACAAATGAAGATCTCA 3720
Db 3661 TATCTTAAATACCCCAACCAAGAGATGAGTTCATTCAATAGGACAAATGAAGATCTCA 3720
Qy 3721 TTCCTTACCACATCACTAGTATTTGGTTAAAAATTTTATATAGTTTTCAGACAATCGTT 3780
Db 3721 TTCCTTACCACATCACTAGTATTTGGTTAAAAATTTTATATAGTTTTCAGACAATCGTT 3780
Qy 3781 GCTAATCTTATCTTGCATTTTGTATGTTCTGTCTGTATTCCTTATATAGCACCTCAG 3840
Db 3781 GCTAATCTTATCTTGCATTTTGTATGTTCTGTCTGTATTCCTTATATAGCACCTCAG 3840
Qy 3841 GCAAGTAGAGTGGCTGGAAGTTTGTGAGTTTGTGGAGTGGAGTTCACAAATATC 3900
Db 3841 GCAAGTAGAGTGGCTGGAAGTTTGTGAGTTTGTGGAGTGGAGTTCACAAATATC 3900
Qy 3901 TATAGTTGATCGAACTACAATPAGCAACATGTGTCGGAATATGTTGCTATCCTCAGCTT 3960
Db 3901 TATAGTTGATCGAACTACAATPAGCAACATGTGTCGGAATATGTTGCTATCCTCAGCTT 3960
Qy 3961 TTTCCCTGTTGACAAATGACATTAACAACTTTAGAACATACACAGGTAAAGAAAGA 4020
Db 3961 TTTCCCTGTTGACAAATGACATTAACAACTTTAGAACATACACAGGTAAAGAAAGA 4020
Qy 4021 TCACGTAGAAATAACATGTTACATTTCCAAATGTTTGCATTAATTTTATAAATTAATACC 4080
Db 4021 TCACGTAGAAATAACATGTTACATTTCCAAATGTTTGCATTAATTTTATAAATTAATACC 4080
Qy 4081 TTATCCATGTTATTTACTACTACAAAATACATTTATGTTGAAACAAACACTTCAAGCA 4140
Db 4081 TTATCCATGTTATTTACTACTACAAAATACATTTATGTTGAAACAAACACTTCAAGCA 4140
Qy 4141 AACATCAGATCTCTTAAAGAGCTGTGTGCTCAAAACCCCTAGTTCCTGTGACACATTG 4200
Db 4141 AACATCAGATCTCTTAAAGAGCTGTGTGCTCAAAACCCCTAGTTCCTGTGACACATTG 4200
Qy 4201 AAAGCAATTTAAAGGAATTTTCAACCAATGATTCCTGACTGTTGCTTCCCAATAATGA 4260
Db 4201 AAAGCAATTTAAAGGAATTTTCAACCAATGATTCCTGACTGTTGCTTCCCAATAATGA 4260
Qy 4261 TGGATACCTCCCTCTACTTTAGGGGTACATAGGTGCAATTTAATGAGTACGAGCCCTTAAA 4320
Db 4261 TGGATACCTCCCTCTACTTTAGGGGTACATAGGTGCAATTTAATGAGTACGAGCCCTTAAA 4320
Qy 4321 CATATTACAGCAGTCCCTTCT 4343
Db 4321 CATATTACAGCAGTCCCTTCT 4343

RESULT 2
US-60-243-468-591
; Sequence 591, Application US/60243468
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET PROTEINS,
; FILE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000929
; CURRENT APPLICATION NUMBER: US/60/243,468
```

```
; CURRENT FILING DATE: 2000-10-27
; NUMBER OF SEQ ID NOS: 2121
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 591
; LENGTH: 62724
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(62724)
; OTHER INFORMATION: n = A,T,C or G
US-60-243-468-591

Query Match 74.8%; Score 3248.4; DB 68; Length 62724;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 3427; Conservative 3; Mismatches 34; Indels 20; Gaps 14;

Qy 872 TTCCTGGGTGGGGGTTGGAGGCATTGAAACAGAGCAGTTATGCTTGGTCTGCCAGTTT 931
Db 36056 TTATTACGTTTGGGGTTGGAGGCATTGAAACAGAGCAGTTATGCTTGGTCTGCCAGTTT 36115
Qy 932 CTCCTTACTTTACCAGAGTGGTTGGATGTGAGTTAACTGGGTCAACAAACCCCTTTGTTA 991
Db 36116 CTCCTTACTTTACCAGAGTGGTTGGATGTGAGTTAACTGGGTCAACAAACCCCTTTGTTA 36175
Qy 992 CAT-CCATAGATGTTGTTCTTGGTATTACAAAGGTAAGTTAAAGTTGTGTAGCTCTATG 1050
Db 36176 CATCCCATAGATGTTGTTCTTGGTATTACAAAGGTAAGTTAAAGTTGTGTAGCTCTATG 36235
Qy 1051 ACTTACTGAACATATTTTATAAAAATTTGAAGCTCTATGAGAGCAGGGATTTGGGTT 1110
Db 36236 ACTTACTGAACATATTTTATAAAAATTTGAAGCTCTATGAGAGCAGGGATTTGGGTT 36295
Qy 1111 CATTACTGCATCCTCAGGTCCTTGACGTTAGCCACATCATCATAGTTATCATAGTAATA 1170
Db 36296 CATTACTGCATCCTCAGGTCCTTGACGTTAGCCACATCATCATAGTTATCATAGTAATA 36355
Qy 1171 ACAACAAACAGAGCATTAGTTTGTACTAATAAATACAAAGAAATTTGTGTGTTCACTT 1230
Db 36356 ACAACAAACAGAGCATTAGTTTGTACTAATAAATACAAAGAAATTTGTGTGTTCACTT 36415
Qy 1231 ATGTTAGTCAATTTAGTCCCTTATACAAACCCCTGTGAGATGGATATATCTATCTCAT 1290
Db 36416 ATGTTAGTCAATTTAGTCCCTTATACAAACCCCTGTGAGATGGATATATCTATCTCAT 36475
Qy 1291 TGTAACTCTGAGAAAATAAGGTACAGTAGGTTTAGTGACCTTACCAGAGGTCGAAGGC 1350
Db 36476 TGTAACTCTGAGAAAATAAGGTACAGTAGGTTTAGTGACCTTACCAGAGGTCGAAGGC 36535
Qy 1351 CTGAGTA-TAAGGGGTAGACAAAGATTCCAGGCAGTCCAGATTCCTTGAGTCAATGT-CT 1406
Db 36536 CTGAGTAGTAAGGGGTAGACAAAGATTCCAGGCAGTCCAGATTCCTTGAGTCCATGTCTT 36595
Qy 1407 AACCATATGCTTATAGTGCCTTGTGCCCTTAATAAACACTTGGTGGAGTACATATTT 1466
Db 36596 AACCATATGCTTATAGTGCCTTGTGCCCTTAATAAACACTTGGTGGAGTACATATTT 36655
Qy 1467 TTTTCTCTCTTTTAACTTGAATTTAAAAAATGTTTAAAGCAAAAGTTCGAMTGTGTCGTC 1526
Db 36656 TTTTCTCTCTTTT-ACTTGAATTTAAAAAATGTTTAAAGCAAAAGTTCGAMTGTGTCGTC 36714
Qy 1527 TTTAATTAATTTATTTGCCCGTTAGAAACTGTGCTCTACTAAGTAATGCTTTCAAAAAC 1586
Db 36715 TTTAATTAATTTATTTGCCCATTTAGAAACTGTGCTCTACTAAGTAATGCTTTCAAAAAC 36774
Qy 1587 ATGGACTGTAGAAATGTGATATATCATTTTTCTGTGCCGTTTTAACATTTCTCTGGATT 1646
Db 36775 ATGGACTGTAGAAATGTGATATATCATTTTTCTGTGCCGTTTTAACATTTCTCTGGATT 36834
Qy 1647 ATTATGTAATAATCTCTCTCTCTGAATTTTAAAAACTGGCTTCAAGACTTCAATACATA 1706
Db 36835 ATTATGTAATAATCTCTCTCTCTGAATTTTAAAAACTGGCTTCAAGACTTCAATACATA 36894
```


Db 39048 GTTGTCTGAGTTTTTGGAGTGGAGTTTACAAATATCTATAGTTGATGCAACTACAAT 39107
QY 3922 AGCAACATGTGTCGGGAATATGTTGCTATCTCAGCTTTTTCCTGTTGACAAATGTGAC 3981
Db 39108 AGCAACATGTGTCGGGAATATGTTGCTATCTCAGCTTTTTCCTGTTGACAAATGTGAC 39167
QY 3982 ATTAACAAATTTAGAACATACAGGTAAAGAGATAAAGATCACTAGATAAATCAATGTTAC 4041
Db 39168 ATTAACAAATTTAGAACATACAGGTAAAGAGATAAAGATCACTAGATAAATCAATGTTAC 39227
QY 4042 ATTTCCAAATGTTTGATAATATTTTAAATATCTACCTTATCCATGTTATTTACTACT 4101
Db 39228 ATTTCCAAATGTTTGATAATATTTTAAATATCTACCTTATCCATGTTATTTACTACT 39287
QY 4102 CACAAAATTTACATGTTGAAACACAACTTTCAAGCAACATCAGATGCTTTTAAAGA 4161
Db 39288 CACAAAATTTACATGTTGAAACACAACTTTCAAGCAACATCAGATGCTTTTAAAGA 39347
QY 4162 GTGTTGTGTCCTCAACCTAGTTCCCTGTGACACATTTGAAGCAATTTAAAGGAATAT 4221
Db 39348 GTGTTGTGTCCTGAACCTAGTTCCCTGTGACACATTTGAAGCAATTTAAAGGAATAT 39407
QY 4222 TCAACCAATTTGATCTGACTGACTGTTT-CCATAATGATGGATACCTCCCTCTACT 4280
Db 39408 TCAACCAATTTGATCTGACTGACTGTTTCCCTATAATGATGGATACCTCCCTCTACT 39467
QY 4281 TAGGGGTGATAGG-TGCAATTTAAGGATCAGCCCTTAAACATATTCACAGCAGTCCCC 4339
Db 39468 TAGGGGTGATAGGTTGCAATTTAATGGAAGTCAGCCTTTAAACATATTCACAGCAGTCCCC 39527
QY 4340 TTCT 4343
Db 39528 TTCT 39531

RESULT 3
US-60-243-742-99
; Sequence 99, Application us/60243742
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000933
; CURRENT APPLICATION NUMBER: US/60/243,742
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 378
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 99
; LENGTH: 62724
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(62724)
; OTHER INFORMATION: n = A,T,C or G
US-60-243-742-99

Query Match 74.8%; Score 3248.4; DB 68; Length 62724;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 3427; Conservative 3; Mismatches 34; Indels 20; Gaps 14;
QY 872 TTCCTGGGTGGGGGTTGGAGCATTGAAACAGAGCAGTTATGCTTGGTCTGCCAGTTT 931
Db 36056 TTATTACGTTAGGGTTGGAGGCATTGAAACAGAGCAGTTATGCTTGGTCTGCCAGTTT 36115
QY 932 CTCCTTACTTACAGAGGTGGTGGATGATGAGTTAACTGGGTGTCATCAACCCCTTTTGTTA 991
Db 36116 CTCCTTACTTACAGAGGTGGTGGATGATGAGTTAACTGGGTGTCATCAACCCCTTTTGTTA 36175
QY 992 CAT-CCATAGATGTTCTCTGTTATACAAAGGTAAGTTAGTGGTAGCTCTATG 1050
Db 36176 CATCCCATAGATGTTCTCTGTTATACAAAGGTAAGTTAGTGGTAGCTCTATG 36235

QY 1051 ACTTACTGAACATTATTTTATATAAAATTTGAAGAGCTCTATGAGAGCAGGGATTTGGGTT 1110
Db 36236 ACTTACTGAACATTATTTTATATAAAATTTGAAGAGCTCTATGAGAGCAGGGATTTGGGTT 36295
QY 1111 CATTACTGCATCCTCAGGTCCTCTGACGTTAGCCACATCATATAGTTATCATATGAATA 1170
Db 36296 CATTACTGCATCCTCAGGTCCTCTGACGTTAGCCACATCATATAGTTATCATATGAATA 36355
QY 1171 ACAACAAACAGACATTTAGTTTGTACTATAAATAACAAGAAATTTGTTGTTGTCACATT 1230
Db 36356 ACAACAAACAGACATTTAGTTTGTACTATAAATAACAAGAAATTTGTTGTTGTCACATT 36415
QY 1231 ATGTTAGCTCATTTAGTTCCTTATAACAAGCCTGTGAGATGGATCTATTACTATTCTCAT 1290
Db 36416 ATGTTAGCTCATTTAGTTCCTTATAACAAGCCTGTGAGATGGATCTATTACTATTCTCAT 36475
QY 1291 TGTAACCTGAGAAAACTAAGGTACAGTAGGGTTTGTAGTCTACCAAGGGTCGAGGC 1350
Db 36476 TGTAACCTGAGAAAACTAAGGTACAGTAGGGTTTGTAGTCTACCAAGGGTCGAGGC 36535
QY 1351 CTGAGTA-TAAGGGTAGACAAAGATTCAGGCAGTCAGATTTCTTGTAGTCTCATGT---CT 1406
Db 36536 CTGAGTAGTAAGGGTAGACAAAGATTCAGGCAGTCAGATTTCTTGTAGTCTCATGTCTCT 36595
QY 1407 AACCATTTATGCCCTTATAGTGCCTTGTTCCTTAAATAAACACTTGTGCTGACTACATATTT 1466
Db 36596 AACCATTTATGCCCTTATAGTGCCTTGTTCCTTAAATAAACACTTGTGCTGACTACATATTT 36655
QY 1467 TTTTCTCTTTTTTAACTTGAATTAACAAAAATGTTTACCAAAAGTTGATGTCGTC 1526
Db 36656 TTTTCTCTTTTTT-ACCTGAAATTTAAAAAAATGTTTACCAAAAGTTGATGTCGTC 36714
QY 1527 TTTAAATTAATTTATTTGCCCGTTAGAAACTGTTGCTCTACTAAGTAATGCTTTCAAAAAC 1586
Db 36715 TTTAAATTAATTTATTTGCCCGTTAGAAACTGTTGCTCTACTAAGTAATGCTTTCAAAAAC 36774
QY 1587 ATGAGCTGTAGAAATGTGATATATCATTTTCTGTTGCCGTTTTTAACATTTCTCTGATTT 1646
Db 36775 ATGAGCTGTAGAAATGTGATATATCATTTTCTGTTGCCGTTTTTAACATTTCTCTGATTT 36834
QY 1647 ATATGTAAAAATCTCTCTCTGAAATTTTAAAAATCTGCTTCAAGATTTCAATACATA 1706
Db 36835 ATATGTAAAAATCTCTCTCTGAAATTTTAAAAATCTGCTTCAAGATTTCAATACATA 36894
QY 1707 CACTGAGCTTTGTTAAGCATATTAATACACAGGCTCAGGATTTCTAGTGAACAATAAT 1766
Db 36895 CACTGAGCTTTGTTAAGCATATTAATACACAGGCTCAGGATTTCTAGTGAACAATAAT 36954
QY 1767 TGTAACCTCTCTCTCTCTCTGAAATGCTGCTGCTTGTAACTTTTATTTTAAATTAATCTTA 1826
Db 36955 TGTAACCTCTCTCTCTCTGAAATGCTGCTGCTTGTAACTTTTATTTTAAATTAATCTTA 37014
QY 1827 TTTTGTAAAAATGATGTACCTGGAATAATGTTCCACATATATTTCCCAATTTTGTAGTCCCAAT 1886
Db 37015 TTTTGTAAAAATGATGTACCTGGAATAATGTTCCACATATATTTCCCAATTTTGTAGTCCCAAT 37074
QY 1887 CTCAGCATTTTGGTTAGATTTATTTGATTCACAGTATGATACCAAAAGATCAAGTAGTAATTT 1946
Db 37075 CTCAGCATTTTGGTTAGATTTATTTGATTCACAGTATGATACCAAAAGATCAAGTAGTAATTT 37134
QY 1947 ATGATATGCTCTCTCTCAGCATTTGGTATTTGATCTCTTCCCTTAATTTAGAAAAGAT 2006
Db 37135 ATGATATGCTCTCTCTCAGCATTTGGTATTTGATCTCTTCCCTTAATTTAGAAAAGAT 37194
QY 2007 TTGSCATCTTAGACAAATTTATTTGATTCACAGTATGATACCAAAAGATCAAGTAGTAATTT 2066
Db 37195 TTGSCATCTTAGACAAATTTATTTGATTCACAGTATGATACCAAAAGATCAAGTAGTAATTT 37254
QY 2067 GGGAAATTCAGGATTTATTCCTTAGAGAAAAGGATATCCCATATATGTTTTTACAGAAATC 2126
Db 37255 GGGAAATTCAGGATTTATTCCTTAGAGAAAAGGATATCCCATATATGTTTTTACAGAAATC 37314

Qy	2127	AAT-CTTACTATTAGACATCTGAAAACTAACCGTCTCTTTTAGCTTCTCTAGCTGTT	2185
Db	37315	ATTTACTTTTACTTTTAGACATCTGAAAACTAACCGTCTCTTTTAGCTTCTCTAGCTGTT	37374
	2186	TTTTCTGTGACAATATTACHTGTGTGTTTTTTCACATTTTACGTTTAAATGTTAAAAAATTAA	2245
Db	37375	TTTTCTGTGACAATATTACHTGTGTGTTTTTTCACATTTTACGTTTAAATGTTAAAAAATTAA	37434
Qy	2246	CTATTATATATGTTTACATTTATTGAATATATATTGATTACTTCTTTTGGAGATCCCTGTTTC	2305
Db	37435	CTATTATATATGTTTACATTTATTGAATATATATTGATTACTTCTTTTGGAGATCCCTGTTTC	37493
Qy	2306	CATTTGTGATCCCTATAGGAATAATCCTGTGTATGTTTTTTTTTGATGAGACAGCATTTGGT	2365
Db	37494	CATTTGTGATCCCTATAGGAATAATCCTGTGTATGTTTTTTTTTGATGAGACAGCATTTGGT	37553
Qy	2366	TTGTAAATATCTAATCTGTGTCTTCTTTCATCTCTAAAAAATAAAACCATAGCCGCGCGGG	2425
Db	37554	TTGTAAATATCTAATCTGTGTCTTCTTTCATCTCTAAAAAATAAAACCATAGCCGCGCGGG	37613
Qy	2426	TGGCTCACGCCGTAAATCCCGACACTTTGGGAGCGCAGCGGGTGGATCATGAGTCCAG	2485
Db	37614	TGGCTCACGCCGTAAATCCCGACACTTTGGGAGCGCAGCGGGTGGATCATGAGTCCAG	37673
Qy	2486	GAGATCGAGACACATCCTGGCTAACAAAGGTGAACCCCGTCTCTACTAAAAATACAAAAA	2545
Db	37674	GAGATCGAGACACATCCTGGCTAACAAAGGTGAACCCCGTCTCTACTAAAAATACAAAAA	37733
Qy	2546	TTAGCCGGCGGGTGGCGGGCGCTGTAGTCCAGCTTCTCGGGAGGCTGAGGCAGAG	2605
Db	37734	TTAGCCGGCGGGTGGCGGGCGCTGTAGTCCAGCTTCTCGGGAGGCTGAGGCAGAG	37793
Qy	2606	AATGGGCTGAACCCCGGAGGGGAGCTTGCAGTCAGCCAAGATTGCGCCACTTGCAGTCCG	2665
Db	37794	AATGGGCTGAACCCCGGAGGGGAGCTTGCAGTCAGCCAAGATTGCGCCACTTGCAGTCCG	37853
Qy	2666	CATCCGGCTGGCGGACAGAGCAGACTCTGTCTC-AAAAAATAAAAAAAAAAAAAA	2725
Db	37854	CATCCGGCTGGCGGACAGAGCAGACTCTGTCTC-AAAAAATAAAAAAAAAAAAAA	37912
Qy	2726	ACCATAAATGAGGAACGCATCTTTACACTTTAGGGTTGAGTTTCGTATCTATAAAAAA	2785
Db	37913	ACCTTAAATGAGGAAA--CATCTTTACACTTTAGGGTTGAGTTTCGTATCTATAAAAAA	37970
Qy	2786	GGGTTTGGATTAAAGTGATCCCTGGCACCTATAAAATCTTAGGGCTTAATATTATTATCATAG	2845
Db	37971	GGGTTTGGATTAAAGTGATCCCTGGCACCTATAAAATCTTAGGGCTTAATATTATTATCATAG	38030
Qy	2846	ATCGAGATAGTTTCATTTAGTGGCTCTTATAGTCACCTTTCCTATACCAATCTGAGA	2905
Db	38031	ATCGAGATAGTTTCATTTAGTGGCTCTTATAGTCACCTTTCCTATACCAATCTGAGA	38090
Qy	2906	CCATTTTACAATTTAGAAAAAGACAAATACTGGTTGGGTTTACGTAGTATATAAACA	2965
Db	38091	CCATTTTACAATTTAGAAAAAGACAAATACTGGTTGGGTTTACGTAGTATATAAACA	38150
Qy	2966	AGAAAAATAATTTTGAAGGAATTAAGTTTGAACACACATGTTTAAACAAATCTTACCAAG	3025
Db	38151	AGAAAAATAATTTTGAAGGAATTAAGTTTGAACACACATGTTTAAACAAATCTTACCAAG	38210
Qy	3026	TGGGATTTGCCTGTGATTAAAGATGCTGTAAACATTTGGGCCAGTAGTTTATAATTTGAAA	3085
Db	38211	TGGGATTTGCCTGTGATTAAAGATGCTGTAAACATTTGGGCCAGTAGTTTATAATTTGAAA	38270
Qy	3086	AATGTTTATAGCCAATATATAATTTTTTTTATTTAAATATACAGTTTCATCAGTCTATATAGT	3145
Db	38271	AATGTTTATAGCCAATATATAATTTTTTTTATTTAAATATACAGTTTCATCAGTCTATATAGT	38330
Qy	3146	ATTTTCATTAAGTCTAAGATGCCATCAGTGGTTAGCAAAACACACTCTTTTATGCACTGCT	3205
Db	38331	ATTTTCATTAAGTCTAAGATGCCATCAGTGGTTAGCAAAACACACTCTTTTATGCACTGCT	38390
Qy	3206	AAGAAAGAAATAAAGGGCTGTGTGCAATGGGCTCACACTGTGGGACGCCAAGCAGGACGA	3265

[illegible]

Db 39468 TAGGGTCATAGTTGCAATTTAATGAAGTCAGCCTTAAACATATTACAGCAGTCCCC 39527
QY 4340 TTCT 4343
Db 39528 TTCT 39531

RESULT 4
US-09-654-935A-257
; Sequence 257, Application US/09654935A
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Asundi, Vinod
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Zhiwei
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; FILE OF INVENTION: Polypeptides
; FILE REFERENCE: 793
; CURRENT APPLICATION NUMBER: US/09/654,935A
; CURRENT FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 271
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 257
; LENGTH: 3280
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2892)
US-09-654-935A-257

Query Match 23.6%; Score 1023; DB 25; Length 3280;
Best Local Similarity 100.0%; Pred. No. 9.5e-111;
Matches 1023; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCCCCAAAAGCAGATACGCCCTTTGAGTACCTATTGAAACATTAATAATGACAGT 60
Db 1 ATGGAGCCCCAAAAGCAGATACGCCCTTTGAGTACCTATTGAAACATTAATAATGACAGT 60
QY 61 TCACATAAGAAGTTCTTCATGATATCTAACTTGGCACCAGATGATGCTTGCCTTAC 120
Db 61 TCACATAAGAAGTTCTTCATGATATCTAACTTGGCACCAGATGATGCTTGCCTTAC 120
QY 121 TCAATACGGGTCTTGTGGAAGCTCTGTACCAAAATGTGATGGCTTTTATGAAGAAG 180
Db 121 TCAATACGGGTCTTGTGGAAGCTCTGTACCAAAATGTGATGGCTTTTATGAAGAAG 180
QY 181 GAAGATGTTATGAACATTTAGACATGGAACCAACCAAGCAATGTTCAAGTGCCTTT 240
Db 181 GAAGATGTTATGAACATTTAGACATGGAACCAACCAAGCAATGTTCAAGTGCCTTT 240
QY 241 TTCCCTGCCCGTGTCTTCTTCAAGATTTTACTGGAATACCAAGCAATGTTGATTTGCT 300
Db 241 TTCCCTGCCCGTGTCTTCTTCAAGATTTTACTGGAATACCAAGCAATGTTGATTTGCT 300
QY 301 GCTATGAGGAGGAGTGAAGTCTTGGAGGTGATCTGGAAGTCCATCCCTGCTTGT 360
Db 301 GCTATGAGGAGGAGTGAAGTCTTGGAGGTGATCTGGAAGTCCATCCCTGCTTGT 360
QY 361 CCGACAGATCTTACAGTTGACCATCTTTTACAAATTTGACTTCAGTAAATGTGCAATACAG 420
Db 361 CCGACAGATCTTACAGTTGACCATCTTTTACAAATTTGACTTCAGTAAATGTGCAATACAG 420
QY 421 ATGACACCAAAATCTCGGAGGTGGTGCCTGCAAGAGCAAGCAAGTCTCTCCACCTTAA 480
Db 421 ATGACACCAAAATCTCGGAGGTGGTGCCTGCAAGAGCAAGCAAGTCTCTCCACCTTAA 480

Db 421 AATGACACAAATCCTGGAGGTGGTGCCTGACCTCAGAAAAGAGAAAGCTCTCTCCACTTAA 480
QY 481 GTGAGGCTTAAGAGCTTCCCTGCAGAGGCCAGACTACCTGCGAGGATCTTCTGATTTCT 540
Db 481 GTGAGGCTTAAGAGCTTCCCTGCAGAGGCCAGACTACCTGCGAGGATCTTCTGATTTCT 540
QY 541 GGAGAACTAGGCGGAAACTCAGGAACATTTTCTTCAGATTCAGAAATACACCCATCTCG 600
Db 541 GGAGAACTAGGCGGAAACTCAGGAACATTTTCTTCAGATTCAGAAATACACCCATCTCG 600
QY 601 TGTCTTTTTCATTTGCAACAGTGCCTGAAACAGTGTAAAAAATCAAGAACTA 660
Db 601 TGTCTTTTTCATTTGCAACAGTGCCTGAAACAGTGTAAAAAATCAAGAACTA 660
QY 661 GAATTCGCGAGAAATCGAGAGAGGCTTTCAGTCTTTTAAAGTGGAGTTCAAGAGTTTAAAG 720
Db 661 GAATTCGCGAGAAATCGAGAGAGGCTTTCAGTCTTTTAAAGTGGAGTTCAAGAGTTTAAAG 720
QY 721 AATGTGCGAGTATCCCTCCTGGAACCTGGAATGCTCATCAAAATAGAAATATTG 780
Db 721 AATGTGCGAGTATCCCTCCTGGAACCTGGAATGCTCATCAAAATAGAAATATTG 780
QY 781 TCAAGAGTGGTTTTTGAAGAAAAGACCTCCTCTCCACAGACAGTGTAGTCGCGACAGAT 840
Db 781 TCAAGAGTGGTTTTTGAAGAAAAGACCTCCTCTCCACAGACAGTGTAGTCGCGACAGAT 840
QY 841 TCACACATAAGATGGTGAATTTAGGATTTCTGGGGTGGGGTGGAGGCATTGAA 900
Db 841 TCACACATAAGATGGTGAATTTAGGATTTCTGGGGTGGGGTGGAGGCATTGAA 900
QY 901 ACAGAAGCAGTTATGCTTGGTCTGCCAGTTTCTTACTTTACCAGAGTGTGGATGT 960
Db 901 ACAGAAGCAGTTATGCTTGGTCTGCCAGTTTCTTACTTTACCAGAGTGTGGATGT 960
QY 961 GAGTTAACTGGTCATCAAAACCCCTTTTGTACATCAGATGTTCTTGTGATATACA 1020
Db 961 GAGTTAACTGGTCATCAAAACCCCTTTTGTACATCAGATGTTCTTGTGATATACA 1020
QY 1021 AAG 1023
Db 1021 AAG 1023

RESULT 5
PCT-US01-08656-4031
; Sequence 4031, Application PC/TUS0108656
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: PCT/US01/08656
; CURRENT FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 4031
; LENGTH: 3290
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (1)..(1869)
; OTHER INFORMATION: 97% homologous to Homo sapiens iron-responsive element-
; OTHER INFORMATION: binding protein/iron regulatory protein 2, accession number
; OTHER INFORMATION: M58511.Smith-Waterman Score=3141.
PCT-US01-08656-4031

Query Match 23.6%; Score 1023; DB 1; Length 3290;
Best Local Similarity 100.0%; Pred. No. 9.5e-111;
Matches 1023; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGACGCCCAAGCAGGATAGCCCTTTGAGTACCTTTATTTGAACATTAATGACAGT 60
Db 1 ATGACGCCCAAGCAGGATAGCCCTTTGAGTACCTTTATTTGAACATTAATGACAGT 60
QY 61 TCACATAGAAGTCTTCGATGATATCTAACTTGGCACCAGATGATGTTCTGCCCTTAC 120
Db 61 TCACATAGAAGTCTTCGATGATATCTAACTTGGCACCAGATGATGTTCTGCCCTTAC 120
QY 121 TCATACGGGCTTGTGGAGGCTGTAGGAAATTTGATGCTTTTAAATCAAGAAG 180
Db 121 TCATACGGGCTTGTGGAGGCTGTAGGAAATTTGATGCTTTTAAATCAAGAAG 180
QY 181 GAAGATGTTATGAACATTTTAGACTGGAACCAACCAAGCAATGTTGAAGTCCCTTT 240
Db 181 GAAGATGTTATGAACATTTTAGACTGGAACCAACCAAGCAATGTTGAAGTCCCTTT 240
QY 241 TTCCCTGCCCGTGTCTTCTTCAAGATTTTACTGGAATACCAAGCAATGTTGAAGTCCCTTT 300
Db 241 TTCCCTGCCCGTGTCTTCTTCAAGATTTTACTGGAATACCAAGCAATGTTGAAGTCCCTTT 300
QY 301 GCTATGAGGAGGAGTCAAAACTTTGGAGGTGATCCTGAGAAAGTCCATCCTGCTTGT 360
Db 301 GCTATGAGGAGGAGTCAAAACTTTGGAGGTGATCCTGAGAAAGTCCATCCTGCTTGT 360
QY 361 CCGACAGATCTTACAGTTGACCATTTTACAAATTTGACTTCAGTAATGTCGAATACAG 420
Db 361 CCGACAGATCTTACAGTTGACCATTTTACAAATTTGACTTCAGTAATGTCGAATACAG 420
QY 421 AATGACCAATCTGGAGGTTGCTGACCTGCAGAAAGCAGGAAAGCTCTCCACTTAAA 480
Db 421 AATGACCAATCTGGAGGTTGCTGACCTGCAGAAAGCAGGAAAGCTCTCCACTTAAA 480
QY 481 GTGAGCCTAAGAGCTTCCCTGCAGAGGCCAGACTACCTGCCAGGATCTTGATGATCT 540
Db 481 GTGAGCCTAAGAGCTTCCCTGCAGAGGCCAGACTACCTGCCAGGATCTTGATGATCT 540
QY 541 GGAGAACTAGCGCCGAAACTCAGGAACATTTTCTTCGAGATTGAGATACACCCATCCTG 600
Db 541 GGAGAACTAGCGCCGAAACTCAGGAACATTTTCTTCGAGATTGAGATACACCCATCCTG 600
QY 601 TGTCCTTTTCAATTTGCAACAGTGCCTGAACCTGAAACAGTGTAAAAATCAAGAAGTA 660
Db 601 TGTCCTTTTCAATTTGCAACAGTGCCTGAACCTGAAACAGTGTAAAAATCAAGAAGTA 660
QY 661 GAAATCGGACAGAAATCGAGAGAGGCTTCAGTTTAAAGTGGAGTTTCAAGAGTTTAAAG 720
Db 661 GAAATCGGACAGAAATCGAGAGAGGCTTCAGTTTAAAGTGGAGTTTCAAGAGTTTAAAG 720
QY 721 AATGTGGCAGTATCCCTTCCTGGAACTGGAATGGCTCATCAATAAACTTAGAATATTG 780
Db 721 AATGTGGCAGTATCCCTTCCTGGAACTGGAATGGCTCATCAATAAACTTAGAATATTG 780
QY 781 TCAAGAGTGGTTTGAAGAAAAAGACCTCCTTCCAGACAGTGTAGTGGGACAGAT 840
Db 781 TCAAGAGTGGTTTGAAGAAAAAGACCTCCTTCCAGACAGTGTAGTGGGACAGAT 840
QY 841 TCACACATAACGATGGTGAATGGTTTAGGGATTCTGGGGTGGGGGTGGAGGCAATTGAA 900
Db 841 TCACACATAACGATGGTGAATGGTTTAGGGATTCTGGGGTGGGGGTGGAGGCAATTGAA 900
QY 901 ACAGAGCAGTATGCTTGGTCTGCCAGTTTCTTACTTTTACCAGAGGTGGTTGGATGT 960
Db 901 ACAGAGCAGTATGCTTGGTCTGCCAGTTTCTTACTTTTACCAGAGGTGGTTGGATGT 960
QY 961 GAGTTAACTGGGTGATCAAAACCCCTTTGTTTACATCCATAGATGTTGTTTGGTATTACA 1020
Db 961 GAGTTAACTGGGTGATCAAAACCCCTTTGTTTACATCCATAGATGTTGTTTGGTATTACA 1020
QY 1021 AAG 1023
Db 1021 AAG 1023
```

```
RESULT 6
US-09-359-922-8303
; Sequence 8303, Application US/09359922
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA
; TITLE OF INVENTION: LIBRARIES
; FILE REFERENCE: 20411-752CON1
; CURRENT APPLICATION NUMBER: US/09/359,922
; CURRENT FILING DATE: 1999-07-22
; EARLIER APPLICATION NUMBER: US 09/205,155
; EARLIER FILING DATE: 1998-12-03
; NUMBER OF SEQ ID NOS: 13203
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8303
; LENGTH: 4684
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(4684)
; OTHER INFORMATION: n = A,T,C or G
US-09-359-922-8303

Query Match 23.6%; Score 1023; DB 17; Length 4684;
Best Local Similarity 100.0%; Pred. No. 9,1e-111;
Matches 1023; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACGCCCAAGCAGGATAGCCCTTTGAGTACCTTTATTTGAACATTAATGACAGT 60
Db 13 ATGACGCCCAAGCAGGATAGCCCTTTGAGTACCTTTATTTGAACATTAATGACAGT 72
QY 61 TCACATAGAAGTCTTCGATGATATCTAACTTGGCACCAGATGATGTTCTGCCCTTAC 120
Db 73 TCACATAGAAGTCTTCGATGATATCTAACTTGGCACCAGATGATGTTCTGCCCTTAC 132
QY 121 TCAATACGGGCTTGTGGAGGCTGTGACGAAATTTGATGGCTTTTAAATGAAGAAG 180
Db 133 TCAATACGGGCTTGTGGAGGCTGTGACGAAATTTGATGGCTTTTAAATGAAGAAG 192
QY 181 GAAGATGTTATGAACATTTTATAGACTGGAACCAACCAAGCAATGTTGAAGTGCCTTT 240
Db 193 GAAGATGTTATGAACATTTTATAGACTGGAACCAACCAAGCAATGTTGAAGTGCCTTT 252
QY 241 TTCCCTGCCCGTGTCTTCTTCAAGATTTTACTGGAATACCAAGCAATGTTGATTTGCT 300
Db 253 TTCCCTGCCCGTGTCTTCTTCAAGATTTTACTGGAATACCAAGCAATGTTGATTTGCT 312
QY 301 GCTATGAGGAGGAGGCTGAAACTCTTGGAGGTGATCCTGAGAAAGTCCATCCTGCTTGT 360
Db 313 GCTATGAGGAGGAGGCTGAAACTCTTGGAGGTGATCCTGAGAAAGTCCATCCTGCTTGT 372
QY 361 CCGACAGATCTTACAGTTGACCATTTTACAAATTTGACTTCAGTAAATGTCGAATACAG 420
Db 373 CCGACAGATCTTACAGTTGACCATTTTACAAATTTGACTTCAGTAAATGTCGAATACAG 432
QY 421 AATGCACCAATCCTCGAGGTGGTGACCTGACAAAGCAGGAAAGCTCTCCACTTAAA 480
Db 433 AATGCACCAATCCTCGAGGTGGTGACCTGACAAAGCAGGAAAGCTCTCCACTTAAA 492
QY 481 GTGACGCTTAAAGAGCTTCCCTGCAGAGGCCAGACTACCTGCCGAGGATCTTGTGATTTCT 540
Db 493 GTGACGCTTAAAGAGCTTCCCTGCAGAGGCCAGACTACCTGCCGAGGATCTTGTGATTTCT 552
QY 541 GGAGAACTAGGCCGAAACTCAGGAACATTTTCTTCGAGATTGAGAATACACCCATCCTG 600
Db 553 GGAGAACTAGGCCGAAACTCAGGAACATTTTCTTCGAGATTGAGAATACACCCATCCTG 612
QY 601 TGTCCTTTTCAATTTGCAACAGTGCCTGACCTGAAACAGTGTAAAAATCAAGAAGTA 660
Db 613 TGTCCTTTTCAATTTGCAACAGTGCCTGACCTGAAACAGTGTAAAAATCAAGAAGTA 672
```

```
QY 661 GAATTCGGCAGAAATCGAGAGGCTTCAGTCTTTTAAAGTGAAGTTTAAAG 720
Db 673 GAATTCGGCAGAAATCGAGAGGCTTCAGTCTTTTAAAGTGAAGTTTAAAG 732
QY 721 AATGTCGAGTATCCCTCCCTGGAAGTGAATGCTCATCAAAATTAAGTGAATTTG 780
Db 733 AATGTCGAGTATCCCTCCCTGGAAGTGAATGCTCATCAAAATTAAGTGAATTTG 792
QY 781 TCAAGAGTGGTCTTTTGAAGAAAGACCTCTCTTCCAGACAGTGTAGTCGGCACAGAT 840
Db 793 TCAAGAGTGGTCTTTTGAAGAAAGACCTCTCTTCCAGACAGTGTAGTCGGCACAGAT 852
QY 841 TCACACATAACGATGGTGAATGGTTTAGGATCTCTGGGTTGGGGTGGAGGCAATGAA 900
Db 853 TCACACATAACGATGGTGAATGGTTTAGGATCTCTGGGTTGGGGTGGAGGCAATGAA 912
QY 901 ACAGAAGCAGTATGCTTGGTCTCCAGTCTCTTACTTTACCAGAGGTGGTGGATGT 960
Db 913 ACAGAAGCAGTATGCTTGGTCTCCAGTCTCTTACTTTACCAGAGGTGGTGGATGT 972
QY 961 GAGTTAACTGGGTATCAAAACCTCTTTGTTACATCCATAGATGTTGTTGATTTACA 1020
Db 973 GAGTTAACTGGGTATCAAAACCTCTTTGTTACATCCATAGATGTTGTTGATTTACA 1032
QY 1021 AAG 1023
Db 1033 AAG 1035

RESULT 7
US-09-359-922-8303
; Sequence 8303, Application US/0935922A
; GENERAL INFORMATION:
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Liu, Jin
; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA
; FILE OF INVENTION: LIBRARIES
; CURRENT APPLICATION NUMBER: US/09/359,922A
; EARLIER FILING DATE: 1999-07-22
; EARLIER FILING DATE: 1998-12-03
; EARLIER APPLICATION NUMBER: US 09/034,341
; EARLIER FILING DATE: 1998-02-13
; NUMBER OF SEQ ID NOS: 13203
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8303
; LENGTH: 4684
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(4684)
; OTHER INFORMATION: n = A,T,C or G
US-09-359-922-8303

Query Match 23.6%; Score 1023; DB 17; Length 4684;
Best Local Similarity 100.0%; Pred. No. 9,1e-111;
Matches 1023; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGACGCCCAAGCAGGATACGCCCTTTGAGTACCTTATGAAACATTAATGACAGT 60
Db 13 ATGGACGCCCAAGCAGGATACGCCCTTTGAGTACCTTATGAAACATTAATGACAGT 72
QY 61 TCACATAGAGTCTTCGATGTATCTAAACTTGGCACCAGATGATGCTCGCCTTAC 120
Db 73 TCACATAGAGTCTTCGATGTATCTAAACTTGGCACCAGATGATGCTCGCCTTAC 132
QY 121 TCAATACGGGTCTTGTGGAGCTGCTGTACGAAATTTGATGGCTTTTAAATGAAGAAG 180
Db 133 TCAATACGGGTCTTGTGGAGCTGCTGTACGAAATTTGATGGCTTTTAAATGAAGAAG 192
QY 181 GAAGATGTTATGAACATTTTAGACTGGAAACCAACAAAGCAATGTTGAAGTGCCCTTT 240
```

```
Db 193 GAAGATGTTATGAACATTTTAGACTGGAAACCAACAAAGCAATGTTGAAGTGCCCTTT 252
QY 241 TTCCCTGCCCGTCTTCTTCAAGATTTTACTGGAAATACCAGCAATGCTGATTTTGTCT 300
Db 253 TTCCCTGCCCGTCTTCTTCAAGATTTTACTGGAAATACCAGCAATGCTGATTTTGTCT 312
QY 301 GCTATGAGGAGGACAGTGAACACTCTTGGAGGTGATCTCTGAGAAAGTCCATCTGCTTGT 360
Db 313 GCTATGAGGAGGACAGTGAACACTCTTGGAGGTGATCTCTGAGAAAGTCCATCTGCTTGT 372
QY 361 CCACAGATCTTACAGTGTGACCAATCTTTACAAATTTGACTTCAGTAAATGTCGAATACAG 420
Db 373 CCACAGATCTTACAGTGTGACCAATCTTTACAAATTTGACTTCAGTAAATGTCGAATACAG 432
QY 421 AATGCAACCAAACTCCTCGAGGTGGTGAACCTGCAGAAAGCAGGAAGCTCTCTCCACTTAA 480
Db 433 AATGCAACCAAACTCCTCGAGGTGGTGAACCTGCAGAAAGCAGGAAGCTCTCTCCACTTAA 492
QY 481 GTGACGCTTAAGAGCTTCCCTGSCAGAGGCCAGACTACCTGCCGAGGATCTTTGATTTCT 540
Db 493 GTGACGCTTAAGAGCTTCCCTGSCAGAGGCCAGACTACCTGCCGAGGATCTTTGATTTCT 552
QY 541 GGAAACTAGGCGGAAACTCAGGAACATTTCTTCGAGATTCAGAAATACACCCATCCCTG 600
Db 553 GGAAACTAGGCGGAAACTCAGGAACATTTCTTCGAGATTCAGAAATACACCCATCCCTG 612
QY 601 TGTCTTTTCAATTTGCAACCACTGCTGAACCTGAAACAGTCTTAAAAATCAAGAACTA 660
Db 613 TGTCTTTTCAATTTGCAACCACTGCTGAACCTGAAACAGTCTTAAAAATCAAGAACTA 672
QY 661 GAATTCGGCAGAAATCGAGAGGCTTTCAGTCTTTTAAAGTGGAGTTCAGAGCTTTTAAAG 720
Db 673 GAATTCGGCAGAAATCGAGAGGCTTTCAGTCTTTTAAAGTGGAGTTCAGAGCTTTTAAAG 732
QY 721 AATGTCGAGTATCCCTCCTCGAAGTGAATGGCTCATCAATAACTTGAATTTTGT 780
Db 733 AATGTCGAGTATCCCTCCTCGAAGTGAATGGCTCATCAATAACTTGAATTTTGT 792
QY 781 TCAAGAGTGGTCTTTGAAGAAAGACCTCTCTCCAGACAGTCTAGTCCGACAGAT 840
Db 793 TCAAGAGTGGTCTTTGAAGAAAGACCTCTCTCTCCAGACAGTCTAGTCCGACAGAT 852
QY 841 TCACACATAACGATGGTGAATGGTCTTGGGATTTCTGGGTTGGGGTGGAGGCAATGAA 900
Db 853 TCACACATAACGATGGTGAATGGTCTTGGGATTTCTGGGTTGGGGTGGAGGCAATGAA 912
QY 901 ACAGAAGCAGTATGCTTGGTCTGCGAGTCTCTTACTTTACCAGAGGTGGTGGATGT 960
Db 913 ACAGAAGCAGTATGCTTGGTCTGCGAGTCTCTTACTTTACCAGAGGTGGTGGATGT 972
QY 961 GAGTTAACTGGGTATCAAAACCTCTTTGTTACATCCATAGATGTTGTTGATTTACA 1020
Db 973 GAGTTAACTGGGTATCAAAACCTCTTTGTTACATCCATAGATGTTGTTGATTTACA 1032
QY 1021 AAG 1023
Db 1033 AAG 1035

RESULT 8
US-09-919-002-8303
; Sequence 8303, Application US/09919002
; GENERAL INFORMATION:
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Liu, Jin
; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA
; FILE OF INVENTION: LIBRARIES
; FILE REFERENCE: 20411-752CON1
; CURRENT APPLICATION NUMBER: US/09/919,002
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: APPLICATION NUMBER: 09/359,922
; PRIOR FILING DATE: FILING DATE: 1999-07-22
```

;; PRIOR APPLICATION NUMBER: APPLICATION NUMBER: US 09/034,341
;; PRIOR FILING DATE: FILING DATE: 1998-02-13
;; NUMBER OF SEQ ID NOS: 13203
;; SOFTWARE: FASTSEQ for Windows Version 3.0
;; SEQ ID NO 8303
;; LENGTH: 4684
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; NAME/KEY: misc_feature
;; LOCATION: (1)...(4684)
;; OTHER INFORMATION: n = A,T,C or G
US-09-1919-002-8303

Query Match 23.6%; Score 1023; DB 34; Length 4684;
Best Local Similarity 100.0%; Pred. No. 9.le-111;
Matches 1023; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGACGCCCAAGCAGGATAGCCCTTTGAGTACCTTATTGAACATTAATGACAGT 60
DB 13 ATGGACGCCCAAGCAGGATAGCCCTTTGAGTACCTTATTGAACATTAATGACAGT 72
QY 61 TCACATAGAAGTCTTCGATGATCAAACTTGGCACCAGATGATGTTCTGCCCTTAC 120
DB 73 TCACATAAGAAGTCTTCGATGATCAAACTTGGCACCAGATGATGTTCTGCCCTTAC 132
QY 121 TCAATACGGGTCTTTGGAAGCTGCTGTAGCAATTTGATGGCTTTTAAATCAAGAAG 180
DB 133 TCAATACGGGTCTTTGGAAGCTGCTGTAGCAATTTGATGGCTTTTAAATCAAGAAG 192
QY 181 GAAGATGTTATGAACATTTTAGACTGGAAACCAACCAAGCAATGTTGAAGTGCCCTTT 240
DB 193 GAAGATGTTATGAACATTTTAGACTGGAAACCAACCAAGCAATGTTGAAGTGCCCTTT 252
QY 241 TTCCCTGCCCTGTTCTTCTCAAGATTTTACTGGAAATACCAAGCAATGTTGGATTTGCT 300
DB 253 TTCCCTGCCCTGTTCTTCTCAAGATTTTACTGGAAATACCAAGCAATGTTGGATTTGCT 312
QY 301 GCTATGAGGAGGAGGAGTGAACCTCTTGGAGGTGATCCTCAGAAAGTCCATCTGCTTGT 360
DB 313 GCTATGAGGAGGAGGAGTGAACCTCTTGGAGGTGATCCTCAGAAAGTCCATCTGCTTGT 372
QY 361 CCGACAGATCTTACAGTTGACCAATCTTTACAAATTCAGTTCAGTAAATGCAATACAG 420
DB 373 CCGACAGATCTTACAGTTGACCAATCTTTACAAATTCAGTTCAGTAAATGCAATACAG 432
QY 421 AATGCCACCAATCTGAGGTGAGTCTGACCTCGAGAAACGAGAAAGCTCTCCACTTAA 480
DB 433 AATGCCACCAATCTGAGGTGAGTCTGACCTCGAGAAACGAGAAAGCTCTCCACTTAA 492
QY 481 GTGCAGCTTAAGAAGCTTCCCTGCAGAGGCCAGACTACCTGCCGAGGATCTTCTGATTCT 540
DB 493 GTGCAGCTTAAGAAGCTTCCCTGCAGAGGCCAGACTACCTGCCGAGGATCTTCTGATTCT 552
QY 541 GGAGAACTAGGCCGAAACTCAGGAACATTTTCTTCGAGATTCGAGATACACCCATCCCTG 600
DB 553 GGAGAACTAGGCCGAAACTCAGGAACATTTTCTTCGAGATTCGAGATACACCCATCCCTG 612
QY 601 TGTCTCTTTTATTGCAACAGTCCCTGAACTGAAACAGTGTAAATAATCAAGAAGTA 660
DB 613 TGTCTCTTTTATTGCAACAGTCCCTGAACTGAAACAGTGTAAATAATCAAGAAGTA 672
QY 661 GAATTCGGCAAAATTCAGAGAGGCTTCAGTCTTTTAAAGTGGAGTCAAGAGTTTAAAG 720
DB 673 GAATTCGGCAAAATTCAGAGAGGCTTCAGTCTTTTAAAGTGGAGTCAAGAGTTTAAAG 732
QY 721 AATGTCGAGTGTATCCCTCCTCGAACTGGAATGGCTCATCAAAATAAATAGATATTTG 780
DB 733 AATGTCGAGTGTATCCCTCCTCGAACTGGAATGGCTCATCAATAAATAGATATTTG 792
QY 781 TCAAGAGTGGTTTTTGAAGAAAGACCTCTCTTCCGACAGAGTGTAGTCGACACAGAT 840
DB 793 TCAAGAGTGGTTTTTGAAGAAAGACCTCTCTTCCGACAGAGTGTAGTCGACACAGAT 852

QY 841 TCACACATAACGATGCTGAATGTTTAGGATTCGGGGTGGGGGTTGGAGGCATTGAA 900
DB 853 TCACACATAACGATGCTGAATGTTTAGGATTCGGGGTGGGGGTTGGAGGCATTGAA 912
QY 901 ACAGAAGCAGTTATGCTTGGTCTGCCAGTTTCTCTTACTTTACCAGAGCTGTTGGATCT 960
DB 913 ACAGAAGCAGTTATGCTTGGTCTGCCAGTTTCTCTTACTTTACCAGAGCTGTTGGATCT 972
QY 961 GAGTTAACTGGGTCAATCAAAACCCCTTTTGTATACATCCATAGATGTTGTTCTTGGTATTACA 1020
DB 973 GAGTTAACTGGGTCAATCAAAACCCCTTTTGTATACATCCATAGATGTTGTTCTTGGTATTACA 1032
QY 1021 AAG 1023
DB 1033 AAG 1035

RESULT 9
US-09-471-275-1719
; Sequence 1719, Application US/09471275
; GENERAL INFORMATION:
; APLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Contigs Obtained
; FILE REFERENCE: 782
; CURRENT APPLICATION NUMBER: US 09/471,275
; EARLIER FILING DATE: 1999-12-23
; EARLIER APPLICATION NUMBER: US 09/235,076
; EARLIER FILING DATE: 1999-01-20
; EARLIER APPLICATION NUMBER: US 09/234,611
; EARLIER FILING DATE: 1999-01-22
; EARLIER APPLICATION NUMBER: US 09/240,371
; EARLIER FILING DATE: 1999-01-29
; EARLIER APPLICATION NUMBER: US 09/277,227
; EARLIER FILING DATE: 1999-03-25
; EARLIER APPLICATION NUMBER: US 09/271,490
; EARLIER FILING DATE: 1999-03-18
; EARLIER APPLICATION NUMBER: US 09/293,972
; EARLIER FILING DATE: 1999-04-15
; EARLIER APPLICATION NUMBER: US 09/274,861
; EARLIER FILING DATE: 1999-03-23
; EARLIER APPLICATION NUMBER: US 60/125,453
; EARLIER FILING DATE: 1999-03-19
; EARLIER APPLICATION NUMBER: US 60/126,605
; EARLIER FILING DATE: 1999-03-26
; EARLIER APPLICATION NUMBER: US 09/306,350
; EARLIER FILING DATE: 1999-05-07
; EARLIER APPLICATION NUMBER: US 09/399,720
; EARLIER FILING DATE: 1999-09-21
; EARLIER APPLICATION NUMBER: US 09/404,284
; EARLIER FILING DATE: 1999-09-21
; EARLIER APPLICATION NUMBER:
; NUMBER OF SEQ ID NOS: 10451
; SOFTWARE: pt_CT_genes Version 1.0
; SEQ ID NO 1719
; LENGTH: 4194
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (151)...(3040)
; OTHER INFORMATION: similar to gl897827 in the genepept database release 114,
; OTHER INFORMATION: Run with FASTXY 3.3t00, default parameters
US-09-471-275-1719

Query Match 23.5%; Score 1019.8; DB 18; Length 4194;
Best Local Similarity 99.8%; Pred. No. 2.2e-110;
Matches 1021; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGAGCGCCCAAGCAGGATACGCTTTGAGTACCTTATTGAACATTAATGACAGT 60
DB 1 ATGGAGCGCCCAAGCAGGATACGCTTTGAGTACCTTATTGAACATTAATGACAGT 60

Db 151 ATGACGCCCCCAAAAGCAGGATAGCCCTTTGAGTACCTTATTGAAACATTAATGACAGT 210
Qy 61 TCACATAAGAAAGTTCTTCGATGTATCTAAACTTGGCCACCAAGTATGATGTCTGCCTTAC 120
Db 211 TCACATAAGAAAGTTCTTCGATGTATCTAAACTTGGCCACCAAGTATGATGTCTGCCTTAC 270
Qy 121 TCAATACGGGCTTGTGTGGAAGCTGCTGTAGCAAAATGTGATGCTTTTAAAGAAAG 180
Db 271 TCAATACGGGCTTGTGTGGAAGCTGCTGTAGCAAAATGTGATGCTTTTAAAGAAAG 330
Qy 181 GAAGATGTTATGAACATTTTACATTTAGCTGGAACCAACCAACCAATGTTGAAGTCCCTTT 240
Db 331 GAAGATGTTATGAACATTTTACATTTAGCTGGAACCAACCAACCAATGTTGAAGTCCCTTT 390
Qy 241 TTCCCTGCCGCTGTTCTTCAAGATTTTACTGGAATACCAAGCAATGGTGGATTTTGGCT 300
Db 391 TTCCCTGCCGCTGTTCTTCAAGATTTTACTGGAATACCAAGCAATGGTGGATTTTGGCT 450
Qy 301 GCTATGAGGGAGGAGTGAAGAACTTGTGGAGGTGATCCTGAGAAAGTCCATCCTGCTTGT 360
Db 451 GCTATGAGGGAGGAGTGAAGAACTTGTGGAGGTGATCCTGAGAAAGTCCATCCTGCTTGT 510
Qy 361 CCGACAGATCTTACAGTTGACCATCTTTACAAATTGACTTCAGTAAATGTCGAATACAG 420
Db 511 CCGACAGATCTTACAGTTGACCATCTTTACAAATTGACTTCAGTAAATGTCGAATACAG 570
Qy 421 AATGACCAAAATCCTGGAGGTGGTACCTGCAGAAAGCAGGAAAGCTCTCCCACTTAAA 480
Db 571 AATGACCAAAATCCTGGAGGTGGTACCTGCAGAAAGCAGGAAAGCTCTCCCACTTAAA 630
Qy 481 GTGAGCCTTAAGAAAGCTTCCCTGCGAGGGCCAGACTACCTGCCAGGAGTCTGTGATCT 540
Db 631 GTGAGCCTTAAGAAAGCTTCCCTGCGAGGGCCAGACTACCTGCCAGGAGTCTGTGATCT 690
Qy 541 GAGAACTAGCCGAACTCAGGACATTTCTTCGAGATGAGATACACCCATCCTG 600
Db 691 GAGAACTAGCCGAACTCAGGACATTTCTTCGAGATGAGATACACCCATCCTG 750
Qy 601 TGCTCTTTTCAATTTGCAACACAGTCCCTGAACTGAAACAGTGTAAAAAATCAAGAAAGTA 660
Db 751 TGCTCTTTTCAATTTGCAACACAGTCCCTGAACTGAAACAGTGTAAAAAATCAAGAAAGTA 810
Qy 661 GAATTCGGCAGAAATCGAGAGGGCTTCAGTTTTTAAAGTGGAGTTCAAGAGTTTAAAG 720
Db 811 GAATTCGGCAGAAATCGAGAGGGCTTCAGTTTTTAAAGTGGAGTTCAAGAGTTTAAAG 870
Qy 721 AATGTCGAGTATCCCTCCCTGGAACCTGGAATGGCTCATCAATAAATAGATATTTG 780
Db 871 AATGTCGAGGATCCCTCCCTGGAACCTGGAATGGCTCATCAATAAATAGATATTTG 930
Qy 781 TCAAGAGTGGTTTTTGAAGAAAGACCTCCTCTTCCAGACAGTGTAGTCGGCAGACAGAT 840
Db 931 TCAAGAGTGGTTTTTGAAGAAAGACCTCCTCTTCCAGACAGTGTAGTCGGCAGACAGAT 990
Qy 841 TCACACATAACGATGGTGAATGGTTTAGGATTTCTGGGGTGGGGTGGAGGCAATTGAA 900
Db 991 TCACACATAACGATGGTGAATGGTTTAGGATTTCTGGGGTGGGGTGGAGGCAATTGAA 1050
Qy 901 ACAGAGCAGTTATGCTTGGTGTGCTGCTCTTCTTACTTTACCAGAGTGGTGGATGT 960
Db 1051 ACAGAGCAGTTATGCTTGGTGTGCTGCTCTTCTTACTTTACCAGAGTGGTGGATGT 1110
Qy 961 GAGTTAACTGGGTATCAAAACCTTTTGTATCCATAGATGTTGTTCTTGGTATTACA 1020
Db 1111 GAGTTAACTGGGTATCAAAACCTTTTGTATCCATAGATGTTGTTCTTGGTATTACA 1170
Qy 1021 AAG 1023
Db 1171 AAG 1173

RESULT 10
US-09-496-914A-7145

; Sequence 7145, Application US/09496914A
; GENERAL INFORMATION:
; APPLICANT: Tang, Yuanhua T.
; APPLICANT: Tillinghast, John
; APPLICANT: Sinku, Ankura
; APPLICANT: Liu, Chenghua
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Contigs Obtained
; FILE REFERENCE: 787
; CURRENT APPLICATION NUMBER: US/09/496,914A
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: US 09/353,690
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 09/034,341
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 09/045,400
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: US 09/321,214
; PRIOR FILING DATE: 1999-05-26
; PRIOR APPLICATION NUMBER: US 09/131,598
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: US 09/431,517
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: US 09/328,351
; PRIOR FILING DATE: 1999-06-04
; PRIOR APPLICATION NUMBER: US 09/332,782
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/235,076
; PRIOR FILING DATE: 1999-01-20
; PRIOR APPLICATION NUMBER: US 09/234,611
; PRIOR FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: US 09/346,956
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: US 09/362,510
; PRIOR FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: US 09/240,371
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: US 09/248,797
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 09/271,490
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: US 09/293,972
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: US 09/274,861
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/125,453
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: US 60/126,605
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: US 09/306,350
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/399,720
; PRIOR FILING DATE: 1999-09-21
; PRIOR APPLICATION NUMBER: US 09/404,284
; PRIOR FILING DATE: 1999-09-21
; PRIOR APPLICATION NUMBER: US 09/465,877
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: NOT YET ASSIGNED
; PRIOR FILING DATE: 1999-01-19
; NUMBER OF SEQ ID NOS: 10410
; SOFTWARE: pt_ct_genes Version 1.02
; SEQ ID NO 7145
; LENGTH: 4194
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (151)...(3040)
; OTHER INFORMATION: similar to gi897827 in the genepept database release 114,
; OTHER INFORMATION: Run with FASTXY 3.3t00, default parameters
US-09-496-914A-7145

Query Match		23.5%;	Score 1019.8;	DB 18;	Length 4194;
Best Local Similarity		99.8%;	Pred. No. 2.2e-110;		
Matches 1021;		Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
Qy	1	ATGGACGCCCAAGCAGATAGCCCTTTGAGTACCTTATTGAACATTTAAATGACAGT	60		
Db	151	ATGGACGCCCAAGCAGATAGCCCTTTGAGTACCTTATTGAACATTTAAATGACAGT	210		
Qy	61	TCACATAAGAGTTCCTTCGATGATCTAAACTTGGCCACCAAGTATGATGTTCTGCCTTAC	120		
Db	211	TCACATAAGAGTTCCTTCGATGATCTAAACTTGGCCACCAAGTATGATGTTCTGCCTTAC	270		
Qy	121	TCAATACGGGTCTTGTGGAAGCTGCTGTACGAAATTTGTGATGGCTTTTAAATGAAGAAG	180		
Db	271	TCAATACGGGTCTTGTGGAAGCTGCTGTACGAAATTTGTGATGGCTTTTAAATGAAGAAG	330		
Qy	181	GAAGATGTTATGAACATTTTAGACTGGAAACCAACAAAGCAATGTTGAAGTGCCTTT	240		
Db	331	GAAGATGTTATGAACATTTTAGACTGGAAACCAACAAAGCAATGTTGAAGTGCCTTT	390		
Qy	241	TTCCCTGCCCTGTTCTTCTCAAGATTTTACTGGAATACCAGCAATGTTGGATTTTGT	300		
Db	391	TTCCCTGCCCTGTTCTTCTCAAGATTTTACTGGAATACCAGCAATGTTGGATTTTGT	450		
Qy	301	GCTATGAGGGAGGCAGTCAAAACTCTTGGAGGTGATCCTGAGAAAGTCCATCCTGCTGT	360		
Db	451	GCTATGAGGGAGGCAGTCAAAACTCTTGGAGGTGATCCTGAGAAAGTCCATCCTGCTGT	510		
Qy	361	CCGACAGATCTTACAGTTCACAGTTCTTACAATTTGACTTCAGTAAATGCAATACAG	420		
Db	511	CCGACAGATCTTACAGTTCACAGTTCTTACAATTTGACTTCAGTAAATGCAATACAG	570		
Qy	421	AATGCACCAAAATCCTGGAGGTGTCACCTGCGAGAGCCAGACTACCTGCCGAGGATCTTGTATCT	480		
Db	571	AATGCACCAAAATCCTGGAGGTGTCACCTGCGAGAGCCAGACTACCTGCCGAGGATCTTGTATCT	630		
Qy	481	GTGCGAGCTTAAAGAGCTTCCCTGCAGAGGCCAGACTACCTGCCGAGGATCTTGTATCT	540		
Db	631	GTGCGAGCTTAAAGAGCTTCCCTGCAGAGGCCAGACTACCTGCCGAGGATCTTGTATCT	690		
Qy	541	GGAGAACTAGCCCAAACTCAGGAACATTTTCTTCGAGATTTGAGAAATACACCCATCCTG	600		
Db	691	GGAGAACTAGCCCAAACTCAGGAACATTTTCTTCGAGATTTGAGAAATACACCCATCCTG	750		
Qy	601	TGTCCTTTTCAATTTGCAACAGTCCCTGAACCTGAAACAGTGTAAAAAATCAAGAAGTA	660		
Db	751	TGTCCTTTTCAATTTGCAACAGTCCCTGAACCTGAAACAGTGTAAAAAATCAAGAAGTA	810		
Qy	661	GAATTCGGCAGAAATCGAGAGAGCTTTCAGTTTTAAAGTGGAGTTCAAGAGTTTAAAG	720		
Db	811	GAATTCGGCAGAAATCGAGAGAGCTTTCAGTTTTAAAGTGGAGTTCAAGAGTTTAAAG	870		
Qy	721	AATGTCGAGTATCCCTCCTCGAATGGAATGGCTCATCAATAACTTAGAATATTG	780		
Db	871	AATGTCGAGGATATCCCTCCTCGAATGGAATGGCTCATCAATAACTTAGAATATTG	930		
Qy	781	TCAAGATGGTTTTGAAGAAAAGAGCTCCTCTTCCTCCAGACAGTGTAGTGGCAGAGAT	840		
Db	931	TCAAGATGGTTTTGAAGAAAAGAGCTCCTCTTCCTCCAGACAGTGTAGTGGCAGAGAT	990		
Qy	841	TCACACATAACGATGGTCAATGGTTTAAAGGATTTCTGGGGTGGGGGTGGAGGCAATTGAA	900		
Db	991	TCACACATAACGATGGTCAATGGTTTAAAGGATTTCTGGGGTGGGGGTGGAGGCAATTGAA	1050		
Qy	901	ACAGAAGCAGTTATGCTTGGTCTGCCAGTTTCTCTTACTTTACAGAGGTGGTGGATGT	960		
Db	1051	ACAGAAGCAGTTATGCTTGGTCTGCCAGTTTCTCTTACTTTACAGAGGTGGTGGATGT	1110		
Qy	961	GAGTTAACTGGTTCATCAAAACCTTTTGTACATCCATAGATGTTGTTCTTGTATATACA	1020		
Db	1111	GAGTTAACTGGTTCATCAAAACCTTTTGTACATCCATAGATGTTGTTCTTGTATATACA	1170		
Qy	1021	AAG 1023			

Db 1171 AAG 1173

|||

RESULT 11

US-09-560-875A-7145

; Sequence 7145, Application US/09560875A

; GENERAL INFORMATION:

; APPLICANT: Tang, Yuanhua T.

; APPLICANT: Tillinghast, John

; APPLICANT: Sinku, Ankura

; APPLICANT: Liu, Chenghua

; APPLICANT: Drmanac, Radoje T.

; TITLE OF INVENTION: Novel Contigs Obtained

; TITLE OF INVENTION: From Various Libraries

; FILE REFERENCE: 787

; CURRENT APPLICATION NUMBER: US/09/560,875A

; CURRENT FILING DATE: 2000-04-27

; Prior application data removed - consult PALM or file wrapper

; NUMBER OF SEQ ID NOS: 10410

; SOFTWARE: pt_CT_genes Version 1.02

; SEQ ID NO 7145

; LENGTH: 4194

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (151)...(3040)

; OTHER INFORMATION: similar to gi897827 in the genepept database release 114

; OTHER INFORMATION: Run with FASTXY 3.3t00, default parameters

US-09-560-875A-7145

Query Match

Best Local Similarity 23.5%; Score 1019.8; DB 22; Length 4194;

Matches 1021; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1	ATGGACGCCCAAGCAGATAGCCCTTTGAGTACCTTATTGAACATTTAAATGACAGT	60
Db	151	ATGGACGCCCAAGCAGATAGCCCTTTGAGTACCTTATTGAACATTTAAATGACAGT	210
QY	61	TCACATAAGAGTTCCTTCGATGATCTAAACTTTGGCACCAGATGATGTTCTGCCTTAC	120
Db	211	TCACATAAGAGTTCCTTCGATGATCTAAACTTTGGCACCAGATGATGTTCTGCCTTAC	270
QY	121	TCAATACGGGTCTTGTGGAAGCTGCTGTACGAAATTTGTGATGGCTTTTAAATGAAGAAG	180
Db	271	TCAATACGGGTCTTGTGGAAGCTGCTGTACGAAATTTGTGATGGCTTTTAAATGAAGAAG	330
QY	181	GAAGATGTTATGAACATTTTAGACTGGAAACCAACAAAGCAATGTTGAAGTGCCTTT	240
Db	331	GAAGATGTTATGAACATTTTAGACTGGAAACCAACAAAGCAATGTTGAAGTGCCTTT	390
QY	241	TTCCCTGCCCTGTTCTTCTCAAGATTTTACTGGAATACCAGCAATGTTGGATTTTGT	300
Db	391	TTCCCTGCCCTGTTCTTCTCAAGATTTTACTGGAATACCAGCAATGTTGGATTTTGT	450
QY	301	GCTATGAGGGAGGCAGTGAACACTCTTGGAGGTGATCCTGAGAAAGTCCATCCTGCTTGT	360
Db	451	GCTATGAGGGAGGCAGTGAACACTCTTGGAGGTGATCCTGAGAAAGTCCATCCTGCTTGT	510
QY	361	CCGACAGATCTTACAGTTGACCATCTTTACAAATTTGACTTCAGTAAATGCAATACAG	420
Db	511	CCGACAGATCTTACAGTTGACCATCTTTACAAATTTGACTTCAGTAAATGCAATACAG	570
QY	421	AATGCACCAAAATCCTGGAGGTGTCACCTGCGAGAGCCAGAGCTCTCTCCACTTAAA	480
Db	571	AATGCACCAAAATCCTGGAGGTGTCACCTGCGAGAGCCAGAGCTCTCTCCACTTAAA	630
QY	481	GTGCGAGCTTAAAGAGCTTCCCTGCAGAGGCCAGACTACCTGCCGAGGATCTTGTGATCT	540
Db	631	GTGCGAGCTTAAAGAGCTTCCCTGCAGAGGCCAGACTACCTGCCGAGGATCTTGTGATCT	690
QY	541	GGAGAACTAGCCCAAACTCAGGAACATTTTCTTCGAGATTTGAGAAATACACCCATCCTG	600

Db	691	GGAGAACTAGCGCGAAACTCAGAAACATTTCTTCGCGAGATTGAGAAATACACCACTCCTG	750
Qy	601	TGTCCCTTTTCATTTCGACACAGTGCCTGAACCTGAACCTGAACAGTGTAAATAATCAAGAAGTA	660
Db	751	TGTCCCTTTTCATTTCGACACAGTGCCTGAACCTGAACAGTGTAAATAATCAAGAAGTA	810
Qy	661	GAATTCGGCAGAAATCGAGAGAGGCTTCAGTTTTTAAAGTGGAGTTCGAAGATTTTAAAG	720
Db	811	GAATTCGGCAGAAATCGAGAGAGGCTTCAGTTTTTAAAGTGGAGTTCGAAGATTTTAAAG	870
Qy	721	AATGTGGCAGTATCCCTCCTCGAACTGGAATGGCTCATCAATAAACTTAGAATATTG	780
Db	871	AATGTGGCAGGATCCCTCCTCGAACTGGAATGGCTCATCAATAAACTTAGAATATTG	930
Qy	781	TCAAGAGTGGTTTTTCAAGAAAAAGACCTCCCTCTTCCCGACAGCTGTAGTCGGCACAGAT	840
Db	931	TCAAGAGTGGTTTTTCAAGAAAAAGACCTCCCTCTTCCCGACAGCTGTAGTCGGCACAGAT	990
Qy	841	TCACACATAACGATGTGAATGGTTTAGGAGATCTCGGGTGGGGGTTGGAGGCATTGAA	900
Db	991	TCACACATAACGATGTGAATGGTTTAGGAGATCTCGGGTGGGGGTTGGAGGCATTGAA	1050
Qy	901	ACAGAAGCAGTTATGCTTGGCTGCCAGTTTCTCTTACTTTACCAGAGGTGGTTGGATGT	960
Db	1051	ACAGAAGCAGTTATGCTTGGCTGCCAGTTTCTCTTACTTTACCAGAGGTGGTTGGATGT	1110
Qy	961	GAGTTAACTGGGTCATCAAAACCTTTTGTTCATACCATAGATGTTGTTCTTGGTATTACA	1020
Db	1111	GAGTTAACTGGGTCATCAAAACCTTTTGTTCATACCATAGATGTTGTTCTTGGTATTACA	1170
Qy	1021	AAG 1023	
Db	1171	AAG 1173	
RESULT 12			
US-60-324-185-26664			
; Sequence 26664, Application US/60324185			
; GENERAL INFORMATION:			
; APPLICANT: Morris, MacDonald			
; APPLICANT: Lal, Preeti			
; APPLICANT: Diep, Dinh			
; TITLE OF INVENTION: METHOD FOR THE IDENTIFICATION OF SEQUENCE POLYMORPHISMS			
; TITLE OF INVENTION: POLYMORPHISMS IDENTIFIED THEREBY			
; FILE REFERENCE: GX-0019-1 P			
; CURRENT APPLICATION NUMBER: US/60/324.185			
; CURRENT FILING DATE: 2001-09-21			
; NUMBER OF SEQ ID NOS: 35862			
; SOFTWARE: PERL Program			
; SEQ ID NO 26664			
; LENGTH: 4729			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: misc.feature			
; OTHER INFORMATION: Incyte ID No: 384853.9			
; NAME/KEY: unsure			
; LOCATION: 1279			
; OTHER INFORMATION: a, t, c, g, or other			
US-60-324-185-26664			

Db	191	TCACAT	AAGAAGTCTTTCGATGTATCTAAACTTGGCACCAAGTATGATGTCTGCCCTTAC	250
Qy	121	TCAA	TACGGGTCTTGTGTGGAAGCTGCTGTACGAAATTTGTGATGGCTTTTAAATGAAGAAG	180
Db	251	TCAA	TACGGGTCTTGTGTGGAAGCTGCTGTACGAAATTTGTGATGGCTTTTAAATGAAGAAG	310
Qy	181	GAAGAT	CTTATGAACATTTTGTAGACTGGAAAAACCAACAAAGCAATGTTGAAGTGCCTTT	240
Db	311	GAAGAT	CTTATGAACATTTTGTAGACTGGAAAAACCAACAAAGCAATGTTGAAGTGCCTTT	370
Qy	241	TTCCCT	CCCGTCTCTTCTTCAAGATTTTACTGGAAATACCAGCAATGGTGGATTTTGTCT	300
Db	371	TTCCCT	CCCGTCTCTTCTTCAAGATTTTACTGGAAATACCAGCAATGGTGGATTTTGTCT	430
Qy	301	GCTAT	GAGGAGGCGAGTGAATACTCTTGAGGTGATCTCTGAGAAAGTCCATCTCGTCTGT	360
Db	431	GCTAT	GAGGAGGCGAGTGAATACTCTTGAGGTGATCTCTGAGAAAGTCCATCTCGTCTGT	490
Qy	361	CCGAC	AGATCTTTACAGTTGACCAATCTTTACAAATTTGACTTTCAGTAAATGTGCAATACAG	420
Db	491	CCGAC	AGATCTTTACAGTTGACCAATCTTTACAAATTTGACTTTCAGTAAATGTGCAATACAG	550
Qy	421	AATCC	ACCAATCCTCGAGGTGGTGAACCTGCAGAAAGCAGGAAAGCTCTCTCCACTTAAA	480
Db	551	AATCC	ACCAATCCTCGAGGTGGTGAACCTGCAGAAAGCAGGAAAGCTCTCTCCACTTAAA	610
Qy	481	GTGCA	GCCTAAGAGCTTCCCTGCAGAGGCCAGACTACCTGCCGAGGATCTTTGTGATTCT	540
Db	611	GTGCA	GCCTAAGAGCTTCCCTGCAGAGGCCAGACTACCTGCCGAGGATCTTTGTGATTCT	670
Qy	541	GGAA	CTAGCGCGAACTCAGGAACATTTCTTTCGAGATTGAGAATACACCCATCCTG	600
Db	671	GGAA	CTAGCGCGAACTCAGGAACATTTCTTTCGAGATTGAGAATACACCCATCCTG	730
Qy	601	TGTC	CTTTTCATTTGCAACCAGTGCCTGAACCTGAAACAGTGTAAAAAATCAAGAAAGTA	660
Db	731	TGTC	CTTTTCATTTGCAACCAGTGCCTGAACCTGAAACAGTGTAAAAAATCAAGAAAGTA	790
Qy	661	GAAT	TCGGCAGAAATCGAGAGAGGCTTCAGTTTTTTTAAAGTGGAGTTCGAAGTTTTTAAAG	720
Db	791	GAAT	TCGGCAGAAATCGAGAGAGGCTTCAGTTTTTTTAAAGTGGAGTTCGAAGTTTTTAAAG	850
Qy	721	AATGT	GCAGTATCCCTCCTCGAACTGGAATGGCTCATCAANTAAACTTAGAATATTG	780
Db	851	AATGT	GCAGTATCCCTCCTCGAACTGGAATGGCTCATCAANTAAACTTAGAATATTG	910
Qy	781	TCAAG	TGGTTTTTGAAGAAAAAGACCTCCTCTCCACAGACAGTGTAGTCGGCACAGAT	840
Db	911	TCAAG	TGGTTTTTGAAGAAAAAGACCTCCTCTCCACAGACAGTGTAGTCGGCACAGAT	970
Qy	841	TCACA	CTAACAGTGTGAATGTTTTAGGATTTCTGGGGTGGGGGTTGGAGGCATTTGAA	900
Db	971	TCACA	CTAACAGTGTGAATGTTTTAGGATTTCTGGGGTGGGGGTTGGAGGCATTTGAA	1030
Qy	901	ACAGA	ACAGTATGCTTGGTCTGCCAGTTTTCTCTTACTTTTACCAGAGGTGGTGGATGT	960
Db	1031	ACAGA	ACAGTATGCTTGGTCTGCCAGTTTTCTCTTACTTTTACCAGAGGTGGTGGATGT	1090
Qy	961	GAGT	TAACGGGTATCAAAACCCCTTTGTTTACATCCATAGATGTTGTTCTTGGTATTACA	1020
Db	1091	GAGT	TAACGGGTATCAAAACCCCTTTGTTTACATCCATAGATGTTGTTCTTGGTATTACA	1150
Qy	1021	AAG	1023	
Db	1151	AAG	1153	

RESULT 13
US-60-339-453-290
; Sequence 290, Application US/60339453
; GENERAL INFORMATION:
; APPLICANT: Wang, Y. Tom
; APPLICANT: Wang, Zhiwei

; TITLE OF INVENTION: Novel Nucleic Acids and
; FILE OF INVENTION: Polypeptides
; FILE REFERENCE: 812
; CURRENT APPLICATION NUMBER: US/60/339,453
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: PCT/US00/35017
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: PCT/US01/02623
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: US 09/496,914
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: US 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: PCT/US01/03800
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 09/515,126
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: US 09/577,409
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: PCT/US01/04927
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: US 09/519,705
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: PCT/US01/04941
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: US 09/649,167
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 09/552,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/770,160
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: PCT/US01/08656
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 412
; SOFTWARE: pt_FL_genes Version 6.0
; SEQ ID NO 63
; LENGTH: 5034
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (76)...(3282)
; US-60-339-453-63

Query Match 23.1%; Score 1004; DB 77; Length 5034;
Best Local Similarity 100.0%; Pred. No. 1.5e-108;
Matches 1004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 GATACGCCCTTTCAGTACCTTATTGAACATTAATGACAGTTCACATGAAGTCTTCG 79
DB 410 GATACGCCCTTTCAGTACCTTATTGAACATTAATGACAGTTCACATGAAGTCTTCG 469
QY 80 ATGTATCTAACTTGGCACCAGTATGTTCTGCTTACTCAATACGGGTCTGTTGG 139
DB 470 ATGTATCTAACTTGGCACCAGTATGTTCTGCTTACTCAATACGGGTCTGTTGG 529
QY 140 AAGTGTCTGACGAATTTGTATGCTTTTAAATGAAGAAGAGATGTTATGAACATTT 199
DB 530 AAGTGTCTGACGAATTTGTATGCTTTTAAATGAAGAAGAGATGTTATGAACATTT 589
QY 200 TAGACTGGAAACCAACAAACCAATGTTGAAGTCCCTTTTCCCTGCCCGTGTCTTC 259

Db 590 TAGACTGGAAACCAACAAACCAATGTTGAAGTCCCTTTTCCCTGCCCGTGTCTTC 649
QY 260 TTCAAGATTTTACTGGAATACCAGCAATGTTGGATTTTGTCTGCTATGAGGAGGCACTGA 319
DB 650 TTCAAGATTTTACTGGAATACCAGCAATGTTGGATTTTGTCTGCTATGAGGAGGCACTGA 709
QY 320 AAACCTTTGGAGTGATCCTCGAGAAAGTCCATCCTGCTTGTCCGACAGATCTTACAGTTG 379
DB 710 AAACCTTTGGAGTGATCCTCGAGAAAGTCCATCCTGCTTGTCCGACAGATCTTACAGTTG 769
QY 380 ACCATCTTTTACAAATTCAGTTCAATGTCGAATACAGATGACCAAAATCCTGGAG 439
DB 770 ACCATCTTTTACAAATTCAGTTCAATGTCGAATACAGATGACCAAAATCCTGGAG 829
QY 440 GTGCTGACCTGCAGAAAGCAGGAAAGCTCTCCACCTTAAAGTGCAGCCTTAAAGAAGCTTC 499
DB 830 GTGCTGACCTGCAGAAAGCAGGAAAGCTCTCCACCTTAAAGTGCAGCCTTAAAGAAGCTTC 889
QY 500 CCTGCAGAGGCCAGACTACCTGCCGAGGATCTTGTGATTTCTGGAGAACTAGGCCGAAACT 559
DB 890 CCTGCAGAGGCCAGACTACCTGCCGAGGATCTTGTGATTTCTGGAGAACTAGGCCGAAACT 949
QY 560 CAGGAACATTTTCTTCGAGATTCAGAAATACACCCATCCTGTCTCTTTTCAATTTGCAAC 619
DB 950 CAGGAACATTTTCTTCGAGATTCAGAAATACACCCATCCTGTCTCTTTTCAATTTGCAAC 1009
QY 620 CAGTCCCTGAACCTGAAACAGTGTAAATAATCAAGAAAGTAGAATTCGGCAGAAATCGAG 679
DB 1010 CAGTCCCTGAACCTGAAACAGTGTAAATAATCAAGAAAGTAGAATTCGGCAGAAATCGAG 1069
QY 680 AGAGGCTTCAGTTTAAAGTGGAGTTCAAGAGTTTAAAGAAATTTGAGAGTGGTCTTGAAG 739
DB 1070 AGAGGCTTCAGTTTAAAGTGGAGTTCAAGAGTTTAAAGAAATTTGAGAGTGGTCTTGAAG 1129
QY 740 CTGGAACCTGGAATGGCTCATCAATAAATAGAAATTTGTCAGAGTGGTCTTGAAG 799
DB 1130 CTGGAACCTGGAATGGCTCATCAATAAATAGAAATTTGTCAGAGTGGTCTTGAAG 1189
QY 800 AAAAGACCTCCTCTTCCAGACAGTGTAGTCCGACAGATTCACACATAACGATGGTGA 859
DB 1190 AAAAGACCTCCTCTTCCAGACAGTGTAGTCCGACAGATTCACACATAACGATGGTGA 1249
QY 860 ATGTTTAGGATTCCTGGGTGGGGGTTGGAGGCAATGAAAGAGAGTGTATGCTTG 919
DB 1250 ATGTTTAGGATTCCTGGGTGGGGGTTGGAGGCAATGAAAGAGAGTGTATGCTTG 1309
QY 920 GTCTGCCAGTTTCTTACTTTACCAGAGTGGTGGATGTGAGTTAACTGGGTCTATCAA 979
DB 1310 GTCTGCCAGTTTCTTACTTTACCAGAGTGGTGGATGTGAGTTAACTGGGTCTATCAA 1369
QY 980 ACCCTTTTGTATCATCCATAGATGTTCTTGTGTTATACAAAG 1023
DB 1370 ACCCTTTTGTATCATCCATAGATGTTCTTGTGTTATACAAAG 1413

RESULT 15
US-09-924-396B-17
; Sequence 17. Application US/09924396B
; GENERAL INFORMATION:
; APPLICANT: Kirsch, Wolff
; APPLICANT: Lennart, Anton
; APPLICANT: Kellin, Wayne
; APPLICANT: Kang, Dae-Kyung
; APPLICANT: Levine, Rodney
; APPLICANT: Rouault, Tracey
; TITLE OF INVENTION: IRON-REGULATING PROTEIN-2 (IRP-2) IS
; FILE REFERENCE: LOMAU.140A
; CURRENT APPLICATION NUMBER: US/09/924,396B
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 60/222,863
; PRIOR FILING DATE: 2000-08-04

; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 2867
; TYPE: DNA
; ORGANISM: Homo Sapiens
; US-09-924-396B-17

Query Match 22.8%; Score 990; DB 34; Length 2867;

Best Local Similarity 100.0%; Pred. No. 7e-107;

Matches 990; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	34	TACCTTATTTGAACATTAATGACAGTTCACATAAGAGTTCCTCGATGTATCTAAACTT	93
DB	1	TACCTTATTTGAACATTAATGACAGTTCACATAAGAGTTCCTCGATGTATCTAAACTT	60
QY	94	GGCACCAAGTATGATGTTCTGCTTACTCAATACGGGTCTGTGGAGGTGCTGTACGA	153
DB	61	GGCACCAAGTATGATGTTCTGCTTACTCAATACGGGTCTGTGGAGGTGCTGTACGA	120
QY	154	AATGTGATGGCTTTTAAATGAAGAAGAGATGTTATGAACATTTAGACTGGAAGCC	213
DB	121	AATGTGATGGCTTTTAAATGAAGAAGAGATGTTATGAACATTTAGACTGGAAGCC	180
QY	214	AAACAAAGCAATGTTGAAGTGCCTTTTCCCTGCCGCTTCTCTTCAAGATTTTACT	273
DB	181	AAACAAAGCAATGTTGAAGTGCCTTTTCCCTGCCGCTTCTCTTCAAGATTTTACT	240
QY	274	GAATACCAAGCAATGGTGGATTTTCTGCTATGAGGAGGCGAGTGAAGTCTTTGGAGGT	333
DB	241	GAATACCAAGCAATGGTGGATTTTCTGCTATGAGGAGGCGAGTGAAGTCTTTGGAGGT	300
QY	334	GATCCTGAGAAAGTCCATCCTGCTGTGCCGACAGATCTTACAGTTGACCATTTCTTACAA	393
DB	301	GATCCTGAGAAAGTCCATCCTGCTGTGCCGACAGATCTTACAGTTGACCATTTCTTACAA	360
QY	394	ATTGACTTCAGTAAATGTGCAATACAGAAATGCACCAATCCTGGAGGTGGTGCACG	453
DB	361	ATTGACTTCAGTAAATGTGCAATACAGAAATGCACCAATCCTGGAGGTGGTGCACG	420
QY	454	AAAGCAGGAAGCTCTCCACTTAAAGTGCAGCCTTAAGAGCTTCCCTGCAGAGGCCAG	513
DB	421	AAAGCAGGAAGCTCTCCACTTAAAGTGCAGCCTTAAGAGCTTCCCTGCAGAGGCCAG	480
QY	514	ACTACCTGCCGAGGATCTTGATTTCTGGAGAACTAGGCCGAAACTCAGGAACATTTTCT	573
DB	481	ACTACCTGCCGAGGATCTTGATTTCTGGAGAACTAGGCCGAAACTCAGGAACATTTTCT	540
QY	574	TGCGAGATTGAGAAATACACCCATCCTGTGCTCTTTTCATTTGCAACCAAGTGCCTGAACCT	633
DB	541	TGCGAGATTGAGAAATACACCCATCCTGTGCTCTTTTCATTTGCAACCAAGTGCCTGAACCT	600
QY	634	GAACAGTGTAAAAAATCAGAAGTAGAATTCGCGAGAAATCGAGAGAGGCTTCAGTTT	693
DB	601	GAACAGTGTAAAAAATCAGAAGTAGAATTCGCGAGAAATCGAGAGAGGCTTCAGTTT	660
QY	694	TTTAAGTGGAGTTCAGAGTTTAAAGAAATGTGGCAGTGATCCCTCTCGAACTGGGAATG	753
DB	661	TTTAAGTGGAGTTCAGAGTTTAAAGAAATGTGGCAGTGATCCCTCTCGAACTGGGAATG	720
QY	754	GCTCATCAATAAATCTAGAAATATTGTCAAGAGTGGTTTTTGAAGAAAAAGACCTCCTC	813
DB	721	GCTCATCAATAAATCTAGAAATATTGTCAAGAGTGGTTTTTGAAGAAAAAGACCTCCTC	780
QY	814	TTCCACAGACAGTGTAGTCGGCACAGATTTCACATACAGATGGTGAATGTTTAGGGATT	873
DB	781	TTCCACAGACAGTGTAGTCGGCACAGATTTCACATACAGATGGTGAATGTTTAGGGATT	840
QY	874	CTGGGGTGGGGGTTGGAGGCATTCAAAACAGAAGCAGTTATGCTTGGTCTGCCAGTTTCT	933
DB	841	CTGGGGTGGGGGTTGGAGGCATTCAAAACAGAAGCAGTTATGCTTGGTCTGCCAGTTTCT	900
QY	934	CTTACTTTTACCAGAGGTGGTTGGATGTGAGTTAACTGGGTCATCAAAACCCCTTTTGTTACA	993

Db	901	CTTACTTTACCAGAGGTGGTTGGATGTGAGTTAACTGGTCAATCAAAACCCCTTTTGTTACA	960
QY	994	TCCATAGATGTTGTTCTTGGTATTACAAAG	1023
DB	961	TCCATAGATGTTGTTCTTGGTATTACAAAG	990

Search completed: February 28, 2003, 11:44:35

Job time : 9172 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 28, 2003, 04:09:08 ; Search time 983 Seconds
(without alignments)
7528.540 Million cell updates/sec

Title: US-09-730-559B-7

Perfect score: 4343

Sequence: 1 atggagcgcacaaagcagg.....attcacagcagtcctctct 4343

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3936061 seqs, 852009584 residues

Total number of hits satisfying chosen parameters: 7872122

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_NA_New.*

- 1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq.*
- 2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq.*
- 3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq.*
- 4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq.*
- 5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq.*
- 6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq.*
- 7: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	1023	23.6	2821	5	US-09-724-676-1037
2	1023	23.6	2821	5	US-09-724-676A-1037
3	877.2	20.2	3770	6	US-10-338-044-2158
4	877.2	20.2	3770	7	US-60-436-643-4106
5	654.8	15.1	5601	6	US-10-144-771-5266
6	368	8.5	47684	5	US-09-949-004-699
C 7	356.6	8.2	114693	1	PCT-US02-09734-3
C 8	336.4	7.7	235033	6	US-10-301-844-1
C 9	336.4	7.7	237326	6	US-10-301-844-2
C 10	325.8	7.5	4384464	5	US-09-948-124-7
C 11	317.4	7.3	465237	1	PCT-US02-25201A-1
C 12	315.6	7.3	14277	7	US-60-427-045-313
C 13	313.8	7.2	267482	5	US-09-949-002-659
C 14	313.8	7.2	267505	5	US-09-949-002-783
C 15	312.6	7.2	859705	5	US-09-948-124-60
C 16	310	7.1	177528	7	US-60-449-629-846
C 17	309.8	7.1	52174	5	US-09-949-002-573
C 18	309.8	7.1	52289	5	US-09-949-002-720
C 19	309.2	7.1	88624	1	PCT-US02-36095-1
C 20	309.2	7.1	63824	1	PCT-US02-34679-347
C 21	309.2	7.1	63824	1	PCT-US02-34679-348
C 22	309.2	7.1	63824	6	US-10-282-174-347
C 23	309.2	7.1	63824	6	US-10-282-174-348
C 24	308.8	7.1	128034	1	PCT-US02-34679-186
C 25	308.8	7.1	128034	1	PCT-US02-34679-187
C 26	308.8	7.1	128034	6	US-10-282-174-186

C 27	308.8	7.1	128034	6	US-10-282-174-187	Sequence 187, App
C 28	308.8	7.1	202100	1	PCT-US02-34679-484	Sequence 484, App
C 29	308.8	7.1	202100	6	US-10-282-174-484	Sequence 484, App
C 30	308.6	7.1	260209	6	US-10-025-966A-23	Sequence 23, Appl
C 31	308.6	7.1	260209	6	US-10-265-071-23	Sequence 23, Appl
C 32	308	7.1	202100	1	PCT-US02-34679-484	Sequence 484, App
C 33	308	7.1	202100	6	US-10-282-174-484	Sequence 484, App
C 34	307.6	7.1	276820	1	PCT-US02-32700-9	Sequence 9, Appl
C 35	307.6	7.1	276820	6	US-10-271-416-9	Sequence 9, Appl
C 36	307	7.1	110096	7	US-60-423-586-68	Sequence 68, Appl
C 37	307	7.1	110096	7	US-60-427-194-68	Sequence 68, Appl
C 38	306.4	7.1	61108	5	US-09-949-002-608	Sequence 608, App
C 39	306.2	7.1	35357	5	US-09-948-124-120	Sequence 3, Appl
C 40	306	7.0	457679	5	US-09-948-124-120	Sequence 120, App
C 41	306	7.0	4384454	5	US-09-948-124-7	Sequence 7, Appl
C 42	305.4	7.0	9451	6	US-10-243-475-131	Sequence 131, App
C 43	305.2	7.0	19793	1	PCT-US02-40718-30	Sequence 30, Appl
C 44	305	7.0	137870	6	US-10-351-951-1	Sequence 1, Appl
C 45	304.8	7.0	601	5	US-09-949-004-717	Sequence 717, App

ALIGNMENTS

RESULT 1
US-09-724-676-1037
; Sequence 1037, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: CompuGen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 125181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1037
; LENGTH: 2821
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-1037

Query Match 23.6%; Score 1023; DB 5; Length 2821;
Best Local Similarity 100.0%; Pred. No. 1.6e-114;
Matches 1023; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGGAGCGCCCAAGCAGGATACGCCCTTGGAGTACCTTATTGAAACATTAAATGACAGT	60
DB	13	ATGGAGCGCCCAAGCAGGATACGCCCTTGGAGTACCTTATTGAAACATTAAATGACAGT	72
QY	61	TCACATAAGAGTCTTCGATGCTATCTAACTTGCACCAAGATGATGTTGCGCTTAC	120
DB	73	TCACATAAGAGTCTTCGATGCTATCTAACTTGCACCAAGATGATGTTGCGCTTAC	132
QY	121	TCAATACGGGCTTCTTGGAAAGCTGCTAGCAAAATGTGATGGCTTTTAAATGAAGAAG	180
DB	133	TCAATACGGGCTTCTTGGAAAGCTGCTAGCAAAATGTGATGGCTTTTAAATGAAGAAG	192
QY	181	GAAGATGTTATGACATTTTATAGCTGGAAACCAACAAACCAATGTTGAAGTGCCTTT	240
DB	193	GAAGATGTTATGACATTTTATAGCTGGAAACCAACAAACCAATGTTGAAGTGCCTTT	252
QY	241	TTCCCTGCCCGTGTCTTCTTCAAGATTTTACTGAAATACCAAGCAATGGTGGATTTGCT	300
DB	253	TTCCCTGCCCGTGTCTTCTTCAAGATTTTACTGAAATACCAAGCAATGGTGGATTTGCT	312
QY	301	GCTATGAGGGAGCGAGTGAACACTCTTGGAGGTGATCTGAGAAGTCCATCTGCTTGT	360
DB	313	GCTATGAGGGAGCGAGTGAACACTCTTGGAGGTGATCTGAGAAGTCCATCTGCTTGT	372
QY	361	CCGACAGATCTTACAGTTGACCATTCCTTACAAATGACTTCAGTAATGTCGAATACAG	420
DB	373	CCGACAGATCTTACAGTTGACCATTCCTTACAAATGACTTCAGTAATGTCGAATACAG	432

QY 421 AATGACCAAAATCCTGGAGTGGTGACCTGCAGAAAGCAGGAAAGCTCTCTCCACTTAAA 480
DB 433 AATGACCAAAATCCTGGAGTGGTGACCTGCAGAAAGCAGGAAAGCTCTCTCCACTTAAA 492
QY 481 GTGAGCCTTAAGAAGTCTCCCTGCAGAGGCCAGACTACTGCGCAGGATCTTGATTTCT 540
DB 493 GTGAGCCTTAAGAAGTCTCCCTGCAGAGGCCAGACTACTGCGCAGGATCTTGATTTCT 552
QY 541 GGAGAACTAGGCGGAAATCAGAAATCTTCTTCCGAGATGGAATACACCCATCTG 600
DB 553 GGAGAACTAGGCGGAAATCAGAAATCTTCTTCCGAGATGGAATACACCCATCTG 612
QY 601 TGTCTCTTTTCATTTGCAACCACTGCTGAACTGAAACAGTGTAAAAAATCAAGAAGTA 660
DB 613 TGTCTCTTTTCATTTGCAACCACTGCTGAACTGAAACAGTGTAAAAAATCAAGAAGTA 672
QY 661 GAATTCGGCAGAAATCGAGAGAGGCTTCAGTTTAAAGTGGAGTTCAAGAGTTTAAAG 720
DB 673 GAATTCGGCAGAAATCGAGAGAGGCTTCAGTTTAAAGTGGAGTTCAAGAGTTTAAAG 732
QY 721 AATGTGGCAGTATCCCTCTGAACTGGAATGGCTCATCAATAAATAGAAATATTG 780
DB 733 AATGTGGCAGTATCCCTCTGAACTGGAATGGCTCATCAATAAATAGAAATATTG 792
QY 781 TCAGAGTGGTTTTGAAGAAAGACCTCTCTCCAGACAGTGTAGTCGGCACAGAT 840
DB 793 TCAGAGTGGTTTTGAAGAAAGACCTCTCTCCAGACAGTGTAGTCGGCACAGAT 852
QY 841 TCACACATAACGATGGTAAAGTGTAGGGATCTGGGGTGGGGGTTGGAGGCATTGAA 900
DB 853 TCACACATAACGATGGTAAAGTGTAGGGATCTGGGGTGGGGGTTGGAGGCATTGAA 912
QY 901 ACAGAGCAGTATGCTGCTGCGAGTTCTCTTACTTTTACCAGAGTGGTTGATGT 960
DB 913 ACAGAGCAGTATGCTGCTGCGAGTTCTCTTACTTTTACCAGAGTGGTTGATGT 972
QY 961 GAGTTAACTGGGTCATCAAAACCTTTTGTATACATCCATAGATGTTCTTGGTATTACA 1020
DB 973 GAGTTAACTGGGTCATCAAAACCTTTTGTATACATCCATAGATGTTCTTGGTATTACA 1032
QY 1021 AAG 1023
DB 1033 AAG 1035

RESULT 2

US-09-724-676A-1037
; Sequence 1037, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724, 676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1037
; LENGTH: 2821
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-1037

Query Match 23.6%; Score 1023; DB 5; Length 2821;
Best Local Similarity 100.0%; Pred. No. 1.6e-114;
Matches 1023; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGCCGCCAAAGCAGATACGCCCTTTGAGTACCTTATTGAAACATTAATGACAGT 60
DB 13 ATGAGCCGCCAAAGCAGATACGCCCTTTGAGTACCTTATTGAAACATTAATGACAGT 72
QY 61 TCACATAGAAGTCTTCTCGATGTATCTFAAACTTGGCACCAGATGATGATGTTCTGCCCTTAC 120
DB 73 TCACATAGAAGTCTTCTCGATGTATCTFAAACTTGGCACCAGATGATGATGTTCTGCCCTTAC 132

RESULT 3

US-10-338-044-2158
; Sequence 2158, Application US/10338044
; GENERAL INFORMATION:
; APPLICANT: MENDRICK, Donna
; APPLICANT: PORTER, Mark
; APPLICANT: JOHNSON, Kory

QY 121 TCATATACGGTCTTGTGTGGAAGCTGCTGTACGAAATGTGTATGGCTTTTAAATGAAGAAG 180
DB 133 TCAATACGGCTCTGTGTGGAAGCTGCTGTACGAAATGTGTATGGCTTTTAAATGAAGAAG 192
QY 181 GAGATGTTATGAAATCTTTAGACTGGAAACCAACAAAGCAATCTTGAAGTGCCCTTT 240
DB 193 GAGATGTTATGAAATCTTTAGACTGGAAACCAACAAAGCAATCTTGAAGTGCCCTTT 252
QY 241 TTCCCTCCCGTGTCTTCTCAAGATTTTACTGGAATACACCAATGGTGGATTTTGCT 300
DB 253 TTCCCTCCCGTGTCTTCTCAAGATTTTACTGGAATACACCAATGGTGGATTTTGCT 312
QY 301 GCTATGAGGAGGAGTGAAGAACTTTTGGAGGTGATCTCGAGAAAGTCCATCTCTGCTTGT 360
DB 313 GCTATGAGGAGGAGTGAAGAACTTTTGGAGGTGATCTCGAGAAAGTCCATCTCTGCTTGT 372
QY 361 CCGACAGATCTTACAGTTGACCATTTTACAAATTCAGTTCAGTAATCTGCAATACAG 420
DB 373 CCGACAGATCTTACAGTTGACCATTTTACAAATTCAGTTCAGTAATCTGCAATACAG 432
QY 421 AATGCACCAATCTCGAGGTGGTGACCTGCAGAAAGCAGGAAAGCTCTCTCCACTTAAA 480
DB 433 AATGCACCAATCTCGAGGTGGTGACCTGCAGAAAGCAGGAAAGCTCTCTCCACTTAAA 492
QY 481 GTGAGCCTTAAGAAGTCTCCCTGCAGAGGCCAGACTACTTGGCCGAGGATCTTGTGATTTCT 540
DB 493 GTGAGCCTTAAGAAGTCTCCCTGCAGAGGCCAGACTACTTGGCCGAGGATCTTGTGATTTCT 552
QY 541 GGAGAACTAGGCCCAAACTCAGGAACATTTTCTTCGAGATGGAATACACCCATCTG 600
DB 553 GGAGAACTAGGCCCAAACTCAGGAACATTTTCTTCGAGATGGAATACACCCATCTG 612
QY 601 TGTCTTTTCTATTGCAACCACTGCTGAACTGAAACAGTGTAAAAAATCAAGAAGTA 660
DB 613 TGTCTTTTCTATTGCAACCACTGCTGAACTGAAACAGTGTAAAAAATCAAGAAGTA 672
QY 661 GAATTCGGCAGAAATCGAGAGAGGCTTCAGTTTAAAGTGGAGTTTCAAGAGTTTAAAG 720
DB 673 GAATTCGGCAGAAATCGAGAGAGGCTTCAGTTTAAAGTGGAGTTTCAAGAGTTTAAAG 732
QY 721 AATGTGGCAGTATCCCTCTGGAACCTGGAATGGCTCATCAATAAATAGAAATATTG 780
DB 733 AATGTGGCAGTATCCCTCTGGAACCTGGAATGGCTCATCAATAAATAGAAATATTG 792
QY 781 TCAGAGTGGTTTTGAAGAAAGACCTCTCTTCCAGACAGTGTAGTCGGCACAGAT 840
DB 793 TCAGAGTGGTTTTGAAGAAAGACCTCTCTTCCAGACAGTGTAGTCGGCACAGAT 852
QY 841 TCACACATAACGATGGTGAATGGTTTAGGGATCTGGGGTGGGGGTTGGAGGCATTGAA 900
DB 853 TCACACATAACGATGGTGAATGGTTTAGGGATCTGGGGTGGGGGTTGGAGGCATTGAA 912
QY 901 ACAGAGCAGTATGCTGCTGCGAGTTCTCTTACTTTTACCAGAGTGGTTGATGT 960
DB 913 ACAGAGCAGTATGCTGCTGCGAGTTCTCTTACTTTTACCAGAGTGGTTGATGT 972
QY 961 GAGTTAACTGGGTCATCAAAACCTTTTGTATACATCCATAGATGTTCTTGGTATTACA 1020
DB 973 GAGTTAACTGGGTCATCAAAACCTTTTGTATACATCCATAGATGTTCTTGGTATTACA 1032
QY 1021 AAG 1023
DB 1033 AAG 1035

APPLICANT: HIGGS, Brandon
APPLICANT: CASTLE, Arthur
APPLICANT: ELASHOFF, Michael
TITLE OF INVENTION: Molecular Cardiototoxicology Modeling
FILE REFERENCE: 44921-5090-01-US
CURRENT APPLICATION NUMBER: US/10/338,044
CURRENT FILING DATE: 2003-01-27
PRIOR APPLICATION NUMBER: US 60/303,819
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US 60/305,623
PRIOR FILING DATE: 2001-07-17
PRIOR APPLICATION NUMBER: US 60/369,351
PRIOR FILING DATE: 2002-04-03
PRIOR APPLICATION NUMBER: US 60/377,611
PRIOR FILING DATE: 2002-05-06
NUMBER OF SEQ ID NOS: 2696
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2158
LENGTH: 3770
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
OTHER INFORMATION: Genbank Accession No. NM_022863
US-10-338-044-2158

Query Match 20.2%; Score 877.2; DB 6; Length 3770;
Best Local Similarity 91.6%; Pred. No. 3.7e-97;
Matches 940; Conservative 0; Mismatches 83; Indels 3; Gaps 1;

QY 1 ATGGACGCCCAAGCAGGATACGCTTTGAGTACCTTATTGAAACATTAATGACAGT 60
DB 12 ATGGACTCCCAAGTCAGGATACACCTTTGAGTACCTTATTGAAACATTAATGAGCAGT 71
QY 61 *TCACATAGAGTTCTTCGATGTATCTAAACTT---GGCACCAGTATGATGTTCTGCCT 117
DB 72 TCACAGAAGAAGTCTTCAATGTACCTAACTTGGAGGCCACCAAGTATGATTTCTGCCT 131
QY 118 TACTCAATPACGGGCTCTTGTGGAAGCTGCTGTACGAATTTGTATGCTTTTAAATGAAG 177
DB 132 TACTCAATPACGGGCTCTTGTGGAAGCTGCTGTACGAATTTGTATGATTTTAAATGAAG 191
QY 178 AAGGAAGATGTTATGACATTTTACAGTGGAAAACCAACCAAGCAATGTTGAAGTCCCC 237
DB 192 AAGGAAGATGTTATATAATATTTTGGACTGGAAAACCAACCAAGCAATGTTGAAGTCCCC 251
QY 238 TTTTTCCTGCCCGTCTTCTTCAAGATTTTACTGGAATACCAGCAATGGTGGATTTT 297
DB 252 TTTTTCCTGCCCGTCTTCTTCAAGATTTTACTGGAATACCAGCAATGGTGGATTTT 311
QY 298 GCTGCTATGAGGAGGAGTGAACACTCTTGGAGTGATCTGTAGAAAGTCCATCTCTGCT 357
DB 312 GCTGCTATGAGGAGGAGTGAACACTCTTGGAGTGATCTGTAGAAAGTCCATCTCTGCT 371
QY 358 TGTCGACAGATCTTACAGTTGACCATCTTTACAAATTTGACTTCAGTAAATGTGCAATA 417
DB 372 TGTCGACAGATCTTACAGTTGACCATCTTTTACAGATTTGACTTCAGTAAATGTGCAATA 431
QY 418 CAGATGCACCAATCTCTGGAGTGGTGACCTGCGAAGAGCAGGAAGCTCTCTCCACTT 477
DB 432 CAGATGCACCAATCTCTGGAGTGGTGACCTGCGAAGAGCAGGAAGCTCTCTCCACTT 491
QY 478 AAGTGCAGCCTTAAGAAGTTCCTCTGCGAGGCCAGACTACCTCCCGAGGATCTGTGAT 537
DB 492 AAGTGCAGCCTTAAGAAGTTCCTCTGCGAGGCCAGACTACCTCCCGAGGATCTGTGAT 551
QY 538 TCTGGAGAACTAGCCCGAAATCTAGGAACATTTTCTTCGCAATTTGAGATACACCCATC 597
DB 552 TCTGGAGAACTAGCCCGAAATCTAGGAACATTTTCTTCGCAATTTGAGATACACCCATC 611
QY 598 CTGTGCTCTTTTCAATTTGCAACAGTGCCTGAACTGAAACAGTGTAAAAAATCAAGAA 657
DB 612 CTGTGCTCTTTTCAATTTGCAACAGTGCCTGAACTGAAACAGTGTAAAAAATCAAGAA 671

QY 658 GTAGAATTGGCAGAAATCGAGAGGCTTCAGTTTTTTTAAAGTGGAGTTCAAGAGTTTTTA 717
DB 672 GTAGAATTGGCAGAAATCGAGAGGCTTCAGTTTTTTTCAAGTGGAGCTCAGAGCTTTT 731
QY 718 AAGAATGTGGCAGTGATCCCTCTCGAACTGGAATGGCTCATCAATAAAGTCTTGAATAT 777
DB 732 AAGAATGTGGCAGTGATCCCTCTCGAACTGGAATGGCTCATCAAGTGAATTTAGAATAT 791
QY 778 TTGTCAGAGTGGTTTTTGAAGAAAAGACCTCTCTTCCAGACAGTGTATGTCGGCACA 837
DB 792 TTGTCAGAGTGGTTTTTGAAGAAAAGACCTCTCTTCCAGACAGTGTATTTGGCACA 851
QY 838 GATTACACATAACGATGTTAGGATGTTAGGATGTTGGGTTGGGGTGGAGGCATT 897
DB 852 GATTCTCATATAACCATGTTGGAATGATTTGGGAATTTCTTTGGGTTGGAGGCAAT 911
QY 898 GAAACAGAAGCAGTTATGCTTGGCTTGCAGTTTCTCTTACTTTACAGAGTGGTTGGA 957
DB 912 GAGACAGAGGCAGTTATGCTTGGCTTGCAGTTTACTTACTTTACAGAGTGGTTGGA 971
QY 958 TGTGAGTTAACTGGGTCATCAAAACCTTTTGTTCATCATAGATGTTGTTCTTGGTATT 1017
DB 972 TGTGAGTTAACTGGGTCATCAAAACCTTTTGTTCATCATAGATGTTGTTCTTGGCATT 1031
QY 1018 ACAAG 1023
DB 1032 ACAAG 1037

RESULT 4
US-60-436-643-4106
; Sequence 4106, Application US/60436643
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Hepatotoxicology Modeling
; FILE REFERENCE: 44921-5038-P15
; CURRENT APPLICATION NUMBER: US/60/436,643
; CURRENT FILING DATE: 2002-12-30
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/331,273
; PRIOR FILING DATE: 2001-11-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 4295
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4106
; LENGTH: 3770
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. U20181
US-60-436-643-4106

Query Match	20.2%;	Score 877.2;	DB 7;	Length 3770;	
Best Local Similarity	91.6%;	Pred. No. 3.7e-97;			
Matches 940;	Conservative 0;	Mismatches 83;	Indels 3;	Gaps 1;	
Qy 1	ATGGAGCCCAAAAGCAGGATACGCCCTTTGAGTACCTTATTGAAACATTAAATGACAGT	60			
Db 12	ATGGACTCCCAAGTCAGGATACACCTTTGAGTACCTTATTGAAACATTAAATGGCAGT	71			
Qy 61	TCACATAGAAGTCTTCGATGTATCTAACTT---GGCACCAGTATGATGTTCTGCCT	117			
Db 72	TCACAGAAGAAGTCTTCAATGTACTTAACCTTTGGAGCCACCAAGTATGATTTCTGCCT	131			
Qy 118	TACTCAATACGGGTCTTCTGGAAGCTGCTGTACGAATTTGTGATGGCTTTTAAATGAAG	177			
Db 132	TACTCAATACGGGTCTTATTGGAAGCTGCTGTACGAATTTGTGATGATTTTAAATGAA	191			
Qy 178	AAGGAAGATGTTATGAACATTTTATAGCTGGAACCAACCAACCAAGCAATGTTGAAAGTCC	237			
Db 192	AAGGAAGATGTTATGAATTTTGGACTTGAACCAACCAACCAAGCAATGTTGAAAGTCC	251			
Qy 238	TTTTTCCCTGCCCGTCTTCTCAAGATTTTACTGGAATACCAGCAATGGTGGATTT	297			
Db 252	TTTTTCCCGCCCGTGTGTTCTTCAAGATTTTCACTGGAATACCAGCAATGGTGGATTT	311			
Qy 298	GCTGCTATGAGGAGGCGAGTGAACACTCTTGGAGGTGATCTCGAAGTCCATCTCTGCT	357			
Db 312	GCTGCTATGAGGAGGCGAATGAACACTCTTGGAGGTGATCTTGAAGTCCACCTGCC	371			
Qy 358	TGTCGACAGATCTTACAGTTGACATTTCTTTTACAAATTTGACTTCAGTAAATGCAATA	417			
Db 372	TGTCCAACAGATCTCACAGTTGACACTCTTTTACAGATTTGACTTCAGTAAATGCAATA	431			
Qy 418	CAGATGCACCAATCTGGAGGTGGTGACCTGCAGAAAGCAGGAAGCTCTCCACTT	477			
Db 432	GAGAATGCACCAATCTGGAGGTGGTGACCTGCAGAAAGCAGGAAGCTCTCCACTT	491			
Qy 478	AAAGTGCAGCTTAAGAAGCTTCCCTGCGAGAGGCCAGACTACTCTCCGAGGATCTTGAT	537			
Db 492	AAAGTACAGCTTAAGAAGCTTCCATGTGCGAGGCCAGACTACTCTCCGCGGATCGTGAT	551			
Qy 538	TCTGAGAACTAGGCCGAACTCAGGAACATTTTCTTCGAGATTTGAGATACACCCATC	597			
Db 552	TCTGAGAACTAAGCCGAACTCAGGAACATTTTCTTCGAGATTTGAGATACACCCATC	611			
Qy 598	CTGTCTCTTTTCAATTCGAACAGTGCCTGAACCTGAAACAGTGTAAATAACAAGAA	657			
Db 612	CTGTCTCTTTTCAATTCGAACAGTGCCTGAACCTGAGACGGTGTAAATAACAAGAA	671			
Qy 658	GTAGAATTCGGCAGAAATCGAGAGGCTTCAGTTTTTTAAAGTGGAGTTCAGAGTTTTA	717			
Db 672	GTAGAATTCGGCAGAAATCGAGAGGCTTCAATTTTTTCAAGTGGAGTTCAGAGTTTT	731			
Qy 718	AGAATTTGGCAGATGATCCCTCTCGGAATCGGAATGGCTATCAATTAACCTTAGAATAT	777			
Db 732	AGAATTTGGCAGTATCCCTCTCGGAATCGGAATGGCTATCAAGTGAACCTTAGAATAT	791			
Qy 778	TTGTCAGAGTGGTTTTTGAAGAAAGACCTCTCTTCCAGACAGTGTAGTCGGACA	837			
Db 792	TTGTCAGAGTGGTTTTTGAAGAGCCGACCTGCTCTTCCAGACAGCGTAAATTTGGACA	851			
Qy 838	GATTCACACATTAACATGGTGAATGGTTTTAGGGATTTCTGGGTTGGGGTTTGGAGCAT	897			
Db 852	GATTCATATAACCATGGTGAATGGATTCTTGGGTTGGGGTTTGGAGCAT	911			
Qy 898	GAAACAGACAGTATGCTTGGCTGCCAGTTCTCTTACTTTACAGAGGTGGTGGGA	957			
Db 912	GAGACAGAGGAGTATGCTTGGCTGCCAGTTCTTACTTTTACAGAGGTGGTGGGA	971			
Qy 958	TGTGAGTTAACTGGGTGATCAAAACCTTTTGTACATCCATAGATGTTGTTCTGGTATT	1017			
Db 972	TGTGAGTTAACTGGGTGATCAAAACCTTTTGTACATCCATAGATATTGTCCTAGGCATT	1031			

Qy 1018	ACAAAG 1023				
Db 1032	ACAAAG 1037				
RESULT 5					
US-10-144-771-5266					
; Sequence 5266, Application US/10144771					
; GENERAL INFORMATION:					
; APPLICANT: VENTER, J. Craig					
; TITLE OF INVENTION: HUMAN GENOME DISCOVERY SYSTEM AND USES THEREOF					
; FILE REFERENCE: CL001321					
; CURRENT APPLICATION NUMBER: US/10/144,771					
; CURRENT FILING DATE: 2002-05-15					
; NUMBER OF SEQ ID NOS: 47235					
; SEQ ID NO 5266					
; LENGTH: 5601					
; TYPE: DNA					
; ORGANISM: HUMAN					
US-10-144-771-5266					
Query Match	15.1%;	Score 654.8;	DB 6;	Length 5601;	
Best Local Similarity	77.1%;	Pred. No. 1.1e-70;			
Matches 911;	Conservative 0;	Mismatches 112;	Indels 159;	Gaps 3;	
Qy 1	ATGGAGCCCAAAAGCAGGATACGCCCTTTGAGTACCTTATTGAAACATTAAATGACAGT	60			
Db 95	ATGGACTCCCAAGTCAGGATACACCTTTGAGTACCTTATTGAAACATTAAATGGCAAT	154			
Qy 61	TCACATAGAAGTCTTTCGATGTATCTAAACTTTGGCACCA-----	100			
Db 155	TCACAGAAGAAGTCTTTCATCTTAACTTTGGCGGCCAACAGTATGTGTAGACAT	214			
Qy 101	-----AGTATGATGTTCTGCCTTACTCAATACGG	129			
Db 215	GACTTCTGTTTACGGTAATATTTTATCTTGAATCAGATATTCTGCTTACTCAATACGG	274			
Qy 130	GTCTTGTGGAACTGCTGTACGAAATTTGTGATGGCTTTTAAATGAAGGAAGATGTT	189			
Db 275	GTCTTATGGAAGTCTGCTGTACGAAATTTGTGATGGCTTTTAAATGAAGGAAGATGTT	334			
Qy 190	ATGAACATTTTATGACGTGGAACCAACCAACCAAGCAATTTGAAAGTGCCCTTTTCCCTGCC	249			
Db 335	ATGAACATTTTGGACTTGGAAACCAACCAACCAAGCAATTTGAAAGTGCCCTTTTCCCTGCC	394			
Qy 250	CGTGTCTTCTTCAAGATTTTACTGGAATACCAAGCAATGTTGATGTTTCTGCTATGAGG	309			
Db 395	CGTGTGTTCTTCAAGATTTTCACTGGAAATACCGCAATGGTGGATTTTCTGCTATGAGG	454			
Qy 310	GAGGAGTGAACACTCTTGGAGGTGATCCTTGAGAAAGTCCATCCTTGTCCGACAGAT	369			
Db 455	GAGGAGTGAACACTCTTGGAGGTGATCCTTGAAGAAAGTCCACCTGCTGTCGACAGAT	514			
Qy 370	CTTACAGTTGACCATTTCTTTTACAAATTTGACTTCAGTAAATGTGCAATACAAATGCACCA	429			
Db 515	CTCAGATGGACACACTTTTACAGATTTGACTTCAGTAAATGTGCAATACAGAAATGCACCA	574			
Qy 430	AATCTCGAGGTGGTGAACCTTGCAGAAAGCAGAAAGCTCTCCACTTAAAGTGCAGCT	489			
Db 575	AATCTCGAGGTGGTGAACCTTGCAGAAAGCAGAAAGCTCTCTCCACTTAAAGTGCAGCT	634			
Qy 490	AAGAAGCTTCCCTCGAGAGGCCAGACTTACCTGCGGAGGATCTTTGTGATTTCTGGAGAACTA	549			
Db 635	AAGAAGCTTCCATGTGCGGCCAGACTTACCTGCGGAGGATCGTGTGATCTTGGAGAGCTA	694			
Qy 550	GGCGAAACTCAGAACATTTCTTCGCAGATTTGAGAAATACACCCCATCTGTGCTCTTT	609			
Db 695	AGCGGAACTCAGAACATTTCTTCACAGATTTGAGAAATACACCCCATCTGTGCTCTTT	754			
Qy 610	CATTTCACCACTGCTGAACCTGAAACAGTCT-----TAAA	647			
Db 755	CATTTCACCACTGCTGAGTATGATGTTTCTTGTAAATAGTTTGTAGTCCAGTACT	814			


```
QY 648 AAATCAAGAGTAGAATTCGGCAGAAATCGAGAGAGGCTTCAGTCTTT----- 694
Db 815 TATTATAAACTTGGAGTTGGTCAGATTGAGCAGTTGGCTAAACGTTTGTAGTGTCAATGTTT 874
QY 695 ----- 694
Db 875 ATTTGTATCTTGATTTATGAATACAGAGTTACACAGTGCACATAAATACACTGTGTCTCTTC 934
QY 695 -----TTAAGTGGAGTTTCAAGAGTTTAAAGAAATGTGGCAGTGTATCCCTCCT 741
Db 935 AATCCAAACAAAATACAGTGGAGCTCAGAGCTTTTAAAGAAATGTGGCAGTGTATCCCGCT 994
QY 742 GGAACGTGAAGGCTCATCAATAAATAGAAATTTGCAAGAGTGGTTTTTGAAGAA 801
Db 995 GGAACGTGAAGGCTCATCAAGTGAACTTAGAAATTTATCAAGAGTAGTTTTTGAAGAA 1054
QY 802 AAAGACCTCTCTTCCACAGAGTGTAGTCGCGCAGAGTTCCACATACAGATGCTGAAT 861
Db 1055 ACGGACCTGTCTTCCACAGAGTGTAGTTGGCAGAGATTCTCATATAACCATGCTGAAC 1114
QY 862 GCTTTAGGATTTCTGGGGTGGGGGTTGGAGGCATTGAAACAGAGCAAGTTATGCTTGGT 921
Db 1115 GGATGGGATTTCTGGGTGGGAGTGTGGTGCATTGAGACAGAGGCAGTTATGCTTGGC 1174
QY 922 CTGCCAGTTTCTTACTTTACAGAGGTGGTTGGATGTGAGTTAACTGGGTGCATCAAAC 981
Db 1175 CTGCCAGTTTACTTTACTTTTACAGAGGTGGTTGGATGTGAGTTAACTGGGTGCATCCAAT 1234
QY 982 CTTTTGTGTACATCATAGATGTTGTTCTTGGTATTACAAAG 1023
Db 1235 GCTTTTGTGTACATCATAGATATTGCTCTTGGCATTACAAAG 1276
```

RESULT 6.

US-09-949-004-699

; Sequence 699, Application US/09949004

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN HUMAN G-PROTEIN COUPLED

; TITLE OF INVENTION: RECEPTOR GENES ASSOCIATED WITH DISEASE, METHODS OF DETECTION

; TITLE OF INVENTION: AND USES THEREOF

; FILE REFERENCE: CL000848

; CURRENT APPLICATION NUMBER: US/09/949,004

; CURRENT FILING DATE: 2002-12-20

; PRIOR APPLICATION NUMBER: 60/232,045

; PRIOR FILING DATE: 2000-09-13

; NUMBER OF SEQ ID NOS: 6961

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 699

; LENGTH: 47684

; TYPE: DNA

; ORGANISM: Human

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)...(47684)

; OTHER INFORMATION: n = A,T,C or G

US-09-949-004-699

Query Match 8.5%; Score 368; DB 5; Length 47684;
Best Local Similarity 61.8%; Pred. No. 9.3e-37;
Matches 697; Conservative 1; Mismatches 401; Indels 29; Gaps 6;

```
QY 2392 CATCCCTAAAAATAAAACCATAGGCGGCGGGTGGCTGCAGCGCTGTAATCCAGCACT 2451
Db 38874 CAGACATAGTAGACAAATATCGCCGCGCGGGTGGCTGCAGCGCTGTAATCCAGCACT 38933
QY 2452 TTGGGAGCGCGAGCGGGGTGGATCATGAGTCAAGGATCGAGACCATCTCGGTCAACAA 2511
Db 38934 TTGGGAGCGCGAGCGGGGTGGATCATGAGTCAAGGATCGAGACCATCTCGGTCAACAA 38993
QY 2512 GGTGAACCCCGTCTCTACTATAAATACAAAAATTAGCCGGCGGGTGGCGGCGCT 2571
Db 38994 GGTGAACCCCGTCTCTACTATAAATACAAAAATTAGCCGGCGGGTGGCGGCGCT 39053
```

RESULT 7

PCT-US02-09734-3/c

; Sequence 3, Application PC/TUS0209734

; GENERAL INFORMATION:

; APPLICANT: PE CORPORATION (NY)

; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,

```
QY 2572 GTAGTCCCAGCTTCTCGGAGGCTGAGGCGAGGAGATGGCGTGAACCCCGGAGGCGGAGC 2631
Db 39054 GTAGTCCCAGCTACTCGGAGGCTGAGGCGAGGAGATGGCGGAGGAGCCCGGAGCGGAGC 39113
QY 2632 TTGCAGTGAAGGATTTGGCCACCTGCAGTCCGAGTCCGGCTCGGCGCAGACAGCGAG 2691
Db 39114 TTGCAGTGAAGGATTTGGCCACCTGCAGTCCGAGTCCGGCGGCGGCGCAGACAGCGAG 39173
QY 2692 ACTCTGTCTCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 2745
Db 39174 ACTCTGTCTCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 39233
QY 2746 TCTTTACACTTAGGTTTGGAGTTTCTGTATCTATAAAAAAGGTTTGGATTAAGTATCC 2805
Db 39234 ATCACATATGGAAGCTAAAAAAGTGGTTATCGTGAAATAGAGAGTAGGTTTCATGGTTAC 39293
QY 2806 CTGGCCTTTATAAATGTTAGGCTTAATATTATTCATAGATCGAGATAGTTTCTTCT 2865
Db 39294 CAGAGGCTAGGAAGGAGTGGGGAGGGGGCGAGAAAGAGA-----GGTGGATTCAAGGG 39348
QY 2866 TAGTCGCTCTCTTAGTCACCTCTTCCATATACCAATCTGAGACCATTTTACAAATTTAGAAAA 2925
Db 39349 TACAAACATACAATTAATAGAGGACAAAGTTCTAGTGTCTAATAGCACAGTAGGGTGA 39408
QY 2926 GACAAATAACTGGTTGGTTACTTGTATAGTATATAACCAAGAAAAATAATTTTAGAAG 2985
Db 39409 CTAGCCTTAACAACAGTTTATATATATTTTCAAGAAATACCTAGAGAAAGATTTTGAATG- 39467
QY 2986 AATTAAGTTTGAACACACATGTTAACAAATCTACCAAAAGTGGGATTTGCCTGTGATTA 3045
Db 39468 -----CTCCCAACACAAAGAAATGATTAATTTGAAGTGTAGCTAGCTGATTT--- 39518
QY 3046 AGATGCTGTAACATTTGGCCAGTAGTTATTAATTTCAAAAAATGTTTATAGCCAATATAT 3105
Db 39519 -GATCATTACACATTTGTATGCAATGATCAAAATATTCATTTGCCCCACAATATGTAAA 39577
QY 3106 AATTTTTTATTAATAATACAGTTTCACTAGTCTATTAGTATTTCATTAAGTCTAAGATG 3165
Db 39578 ATTATTATATGAAATTTTAAATCTTCAAGAAAAAATGTATATTATTCTAGAAATTC 39637
QY 3166 CCATCAGTGGTTAGCAACACCACCTGTTTTTATGCATGCTTAAGAAAGAATAAAGGCTGT 3225
Db 39638 TAAGAAATACCTCTGTGGACCACCACTTTAAAAAATACTTAGGCCAGATGCTGTGGCTCAT 39697
QY 3226 GTGCAGTGGCTCACACCTGTGGGACGCCAAGCAGGAGCATCTTGAGGCCAGAGTTTC 3285
Db 39698 GCTTGTAAATCCACACACTTTGGGAGGCTGAGTGGTGGATCATTGGGGTTCAGGAGTTC 39757
QY 3286 AAGACCAACCTTGTCAACATTTGAAGACCTGCTCTACAAAAAATAAAAAAAGTTTAAAT 3345
Db 39758 GAGACCAGCTTGCCAAAATAGTGAACCCCATCACTATTAAAAATAC-----AAAAAT 39811
QY 3346 TAGCTGGGTGGGTGGCAGCATGCTGTAGTTCAGCTACTCTGGAGCTAAGTGGGAGG 3405
Db 39812 TAGCAGGTGTGATGGTGGCGCCTGTAATCCAGCTACTCAGGAGCTCAGGAGGAGAAA 39871
QY 3406 ATTGCTAGAGCCACGGTGTGGAGGCTGCAATGAGTGTGACACACACCTGCGCTCCAG 3465
Db 39872 ATTGCTTGAACCTGGGAGGAGAGGTTGCAATGAGCCGAGATCATGCACTTCACTCCAG 39931
QY 3466 CGTGGGCAACAGTAGAGACCTCTTTTCTAAAGAAAAAGAAAAA 3513
Db 39932 CCT-GGTGACAGAGTGAGACTTTTCTCTTAAAGAAATAAATAAATAA 39978
```

RESULT 7

PCT-US02-09734-3/c

; Sequence 3, Application PC/TUS0209734

; GENERAL INFORMATION:

; APPLICANT: PE CORPORATION (NY)

; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,

QY	3181	AAACACCACCTGTTTATGCACTGCTAAGAAAGAAATAAAGGGCTGTGTGCAGTGGCTCACA	3240
Db	12617	CTCTACAAAAAATACAAAGCTGGGTGCGGGTGCAGTGGTTTACGCCATATATCCACG	12558
QY	3241	CTGTGGGAGCCCAAGGAGGAGCAATCACTTGGAGCCAGAAAGTTCAAGACCAACCTGGTGC	3300
Db	12557	ACTTTTGGACACCAAAAGTGGGAGGATTGCTTAAGCCAGAAAGTTCAAGACCAACCTGGGT	12498
QY	3301	AACATTGTAAGACCCCTGCTCTACAAAAAANAANAAGTTAAAAATTAGCTGGGTGCGGTG	3360
Db	12497	AACATGTCAAAACTTTGTCTCTACAAAAATAAT---TTAAAAATTAGCTGAGCATAGTG	12442
QY	3361	GCACATGCCCTGTAGTTCCAGCTACTCTGGAGGCTAAGTGGGAGGATTGCTAGAGCCACG	3420
Db	12441	GTGCAGCCGTGTAGTCTCACTTGGAGGCTCAGGTGGGAGGATCACTGAGGCCAG	12382
QY	3421	GTGTTGGAAGCTGCAATGAGCTGTGACCACACCACTCGCGTCCAGCGTGGGCAACAGACT	3480
Db	12381	GAGTGCCAGGCTCAGTGAGCTGTGATGGCTCGCTCACTCCACCTAGGCAACAGACT	12322
QY	3481	GAGACCCCTGTTCTAAAGAAAGAAAGAAAAGGGCTG	3519
Db	12321	GAGACCCCTGTCTCAAAAAAATTTAAAAAAGAATCAATG	12283
RESULT 8			
US-10-301-844-1/c			
; Sequence 1, Application US/10301844			
; GENERAL INFORMATION:			
; APPLICANT: Ruddy, David A.			
; Wolff, Roger K.			
; TITLE OF INVENTION: POLYMORPHISMS IN THE REGION OF THE HUMAN			
; HEMOCHROMATOSIS GENE			
; NUMBER OF SEQUENCES: 26			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Pennie & Edmonds, LLP			
; STREET: 1155 Avenue of the Americas			
; CITY: New York			
; STATE: NY			
; COUNTRY: USA			
; ZIP: 10036-2811			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: Windows			
; SOFTWARE: FastSeq for Windows Version 2.0b			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/10/301,844			
; FILING DATE: 20-Nov-2002			
; CLASSIFICATION: <Unknown>			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: US/08/852,495C			
; FILING DATE: 07-MAY-1997			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Poissant, Brian M			
; REGISTRATION NUMBER: 28,462			
; REFERENCE/DOCKET NUMBER: 8907-0057-999			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: 650-493-4935			
; TELEFAX: 650-493-5556			
; TELEX: 66141 PENNIE			
; INFORMATION FOR SEQ ID NO: 1:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 235033 base pairs			
; TYPE: nucleic acid			
; STRANDEDNESS: single			
; TOPOLOGY: linear			
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:			
US-10-301-844-1			

Db 134325	CGGAGGTTGCAGTGCAGCCGAGATTGTGCCACT-----GCATTCAGCCTGGGCCACAG	134277
Qy 2686	AGCGAGACTCTGTTTCAAAAAAAAAAAAAAAAAAACCAATAAATGAGGAAACGCA	2745
Db 134272	AGCAAGACTCCGCTCAGAGAAAAAAGTCACTTAAATTTAAGAGAAAAA	134213
Qy 2746	TCATTACACATTTAGGTTTGAGTTTCGTATCTATAAAAAGGGTTTGGATTAGTGATCC	2805
Db 134212	TGCATTTGCTTTGGGACTTTTAATATTATTAGTCTACAACTACGCCACCATAGAAATCTGC	134153
Qy 2806	CTGGCACTTATAAATGTTAGGGCTTAATATTATTCATAGATCGAGTAGTTTTCATTCT	2865
Db 134152	TGATTAATACGGGTTCTGTTAAATGGAAACATGCATTTTGGGGGAAAAAGAGGGAGT	134093
Qy 2866	TAGTGGCCTCCTTAGTCACACTCTTCCCTATACCAATCTGAGACCAATTTTACAATTTAGAAA	2925
Db 134092	GTTTGTAGTGATTTTGTGTTTTTACACTTGTGTTATAATAAAATTTTAAGCAATCTTGAGGG	134033
Qy 2926	GACAAATAACTGGTTGGGTTACTTGATAGTATAATAACCAAGAAAAATATTTTACAAGG	2985
Db 134032	ANCAATTTATTCTTACTTGTAACTGCACTGCAATAAAGTTATGAGATAAAGTTACAGCTATATCA	133973
Qy 2986	AATTAAGTTTGAACACCATG--TTACAAATTTCTACCAAGTGG-GATTTGGCTGTGAT	3042
Db 133972	CATACAGTTGTAGCTTTATAAATTATGAATCTTAACAGAAATAATATGCTAATATGAT	133913
Qy 3043	TAAAGATGCTGTAACACATTTGGGCCAGTACTTATAATTTGAAAAATGTTTATAGCCAATA	3102
Db 133912	GAAATGCTATAAATACATTAGAATATATTTTAATAAACCAATTCAGAAGAGGACCAATA	133853
Qy 3103	TATAATTTTTTATTTAAATATACAGTTTTCATCAGTCTA-----TTAGTATTTTCATTAA	3155
Db 133852	CCCAATTTCAAAATCATATTAAATGTAATAATTAATAGGCGACCCAAATATTTCTGGAAT	133793
Qy 3156	GTCTAAGATGCCATCACTGGTTAGCAAAACCACTGTTTATGCAGCTGTAAGAAAGAAAT	3215
Db 133792	TCTTTCTAATAAAAAAATAGTGTAAATACAGTCGCTACTGCACAAATCTGAAGAATTTATG	133733
Qy 3216	AAAGGCTGTGTGCAGTGGCTCACACCTGTGGGAGCGCAGGAGGAGCATCACTTTGAGG	3275
Db 133732	CAGCATAAAAAGTGATTTATCCAGCACTTTTGGGAGCCAGTTGGCGAGATCA--TGAGG	133675
Qy 3276	CCAGAAGTTCAAGACCAACCTGGTCAACATTTGTAAGACCTGTCTCTACAAAAAATAAAA	3335
Db 133674	TCAGGAGTTGGAGACCAAGCCTGACCAACATGGTGAACCCCGCTTTTACTAAAAATA--	133618
Qy 3336	AGTTAAAAATTAGCTGGTGGGTGGCGGCACATGGCTGTAGTTCCAGCTACTCTCGAGGCTA	3395
Db 133617	---CAAAAAATTAGCGGGCTTGGTGGTGCACACTGTAATCCCAAGCTGCTCAGAGGCTA	133561
Qy 3396	AGGTGGGAGGATTGCTAGAGCCACGGTGTGGAGCTGCAATGAGCTGTGACCAACACCAC	3455
Db 133560	AGGCAGGAGAACTGCTTGAATCCAGGAGGTGGAAGTTGCGATGAAGCAAGATCAGCCAC	133501
Qy 3456	TGGCGTCCAGCGTGGGCAACAGAGTGAGACCGCTGTTCTTAAAAAGAAAGAAAAA	3513
Db 133500	TGCACCTCAGGCTGGGGGACAGAGTGTGACTCTGACTCAAAAAAATAAATAAATA	133443
RESULT 10		
US-09-948-124-7/c		
; Sequence 7, Application US/09948124		
; GENERAL INFORMATION:		
; APPLICANT: VENTER, J. Craig		
; TITLE OF INVENTION: COLLECTION OF SINGLE NUCLEOTIDE POLYMORPHISM (SNPs) LOCATED		
; TITLE OF INVENTION: Y, METHODS OF DETECTION, AND USES THEREOF		
; FILE REFERENCE: CL001301		
; CURRENT APPLICATION NUMBER: US/09/948,124		
; CURRENT FILING DATE: 2002-12-23		
; NUMBER OF SEQ ID NOS: 183		
; SEQ ID NO 7		
; LENGTH: 4384464		
; TYPE: DNA		

```

; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(4384464)
; OTHER INFORMATION: n = A,T,C or G
US-09-948-124-7

Query Match          7.5%; Score 325.8; DB 5; Length 4384464;
Best Local Similarity 58.4%; Pred. No. 2.6e-32;
Matches 668; Conservative

Qy 2412 TAGGCCGGCGGGTGGCTCACGCCGTATCCAGACACTTTGGGAGGCCGAGCGGGTG 2471
    |||||
Db 957133 TGGCCGGCGGGTGGTTCACGCCGTATCCAGCACTTTGGGAGGCCGAGGTGGTG 957074

Qy 2472 GATCATGAGTTCAGGAGATCGAGACCATCTCGCTAACAGGTGAACCCCGTCTCTACT 2531
    |||||
Db 957073 GATCACGAGTTCAGGAGATCGAGACCATCTCGCTAACACAGTGAACCCCATCTCTACT 957014

Qy 2532 AAAAATACAAAAATTAGCGCGGGTGGCGGGCCCTGTAGTCCCAGCTTCTCGCGGA 2591
    |||||
Db 957013 AAAAATACAAAAATTAGCTGGCGGTGGTGGCGGGCCCTGTAGTCCCAGCTTACGCGGA 956954

Qy 2592 GGCTGAGCGAGGAATGGCGTGAACCCGGGAGCGGAGCTTGCACTGAGCCAAAGATTCC 2651
    |||||
Db 956953 GGCTGAGCGAGGAATGGCGTGAACCCAGGAGCGGAGCTTGCACTGAGCTGAGATCGT 956894

Qy 2652 GCCACTGCAGTCCGCACTCGCGCTGGGCGACAGAGCGAGACTCTGTTCAAAAAATA 2711
    |||||
Db 956893 GCCACT-----GCACTCGAGCTGGGCAACAGAGTGAGACTCTGTCTCAAAAAAATA 956841

Qy 2712 AAAAAA-----A-----AAAAAACCATTAAT 2734
    |||||
Db 956840 AAAAAAAGCTTTTGTGTTTACTTCCACCCAAATTTGTATACCAACATTTATTTATAT 956781

Qy 2735 GAGGAACGCATCTTTACACTTAGGGTTTGAGTTCTGTATCTATATAAAGGTTTGA 2794
    |||||
Db 956780 GCATATAGATCATTTTAAAAATTTATACCAAGTGAAGTACTATACAAAGTTTGATG 956721

Qy 2795 TTAAGTGATCCCTGGCACTTATAAATGTTTAGGCTTAATATTATTATCATGATCGAGGAT 2854
    |||||
Db 956720 TCTTGTGTTTGTCTTCACTTAAAGTATGCTTCGAGAAATTTCTACATTAAGAAGATAG 956661

Qy 2855 AGTTTCATCTTTAGTCGCCCTCTTAGTCACTCTTCCCTATACCAATCTGAGACCATTTTAC 2914
    |||||
Db 956660 ATATATCTCATTTATTTAGCGGTGGCATGTTATCTTAATGACTTTGATGGACCATGACTG 956601

Qy 2915 AATTTAGAAAGACAATAACTGGTTGGGTACTTGTATAGTATAATAAACCAAGAAAAATA 2974
    |||||
Db 956600 ATTTAAGCTGTCAATACTATTTAGGTAATAGTCTAGATTTTACTTAACACGACAAATGCT 956541

Qy 2975 ATTTTAGAAGAAATTAAGTTTGAACCCACATGTTAACAAATTTACCAAGTGGGATTTG 3034
    |||||
Db 956540 GCAATGAATATCATTTGAGCAGAAATTTGAATGCTATTTACTAGTCTATCTGTCATGGAA 956481

Qy 3035 CCTGTGATTAAGATGCTGAACATTTGGCGCAGTAGTTATTAATTTGAANAATGTTTAT 3094
    |||||
Db 956480 ACAATGGAATTTGTTGATGAAGAAATGTGCAAGTTTATTTATTAATAATTTTCAATCAT 956421

Qy 3095 AGCCAATATATAATTTTTTATTTAAATATACAGTTTTCATCAGTCTATTAGTATTTTCATTA 3154
    |||||
Db 956420 TAAGAAAAAGACTGATAATACATAATTTGTTCTATCCCTCTAATATTTTCAAGATTACC 956361

Qy 3155 AGTCTAGATG-CGATCAGTGGTTAGCAACACCACTGTTTTATGCACTGCTTAAGAAGA 3213
    |||||
Db 956360 ATTCTGGCAATGTTGTCAATACACCCCTTATTTATTAATAATTTGAACCTCTTGGCTGGCC 956301

Qy 3214 ATAAAGGCTGTGTGCAGTGGCTCACACCTGTGGGAGCCCAAGCGAGGAGCATCACTTGA 3273
    |||||
Db 956300 ACAGTGGCTCAGGCCTGTAATCCAGCACTTTGGAGGCTTGAGGAGGTGATCACTGA 956241

Qy 3274 GGCCAGAAGTTTCAAGACCAACCTGGTCAACATTTGAAGACCCCTGTCTCTTCAAAAAAATA 3333
    |||||

```

Db 956240 GGTTCAGGAGTTCAAGACCATCTTGGCCAACTGGTGA-AACCGGCTTACAAAAA 956182
 QY 3334 AAGTTAAATAATAGCTGGTGGTGGGCGGACATGCTGTAGTTCAGTACTCTGGAGGC 3393
 Db 956181 AGAAAAAATACAGACGAGCTGGCAGCTGGCCACTGTAAATCCAGACTACTTGGGCAGC 956122
 QY 3394 TAAGGTGGGAGGATTGCTAGAGCCACCGGTGTGGAAAGCTGCAATGAGCTGTGACACACACC 3453
 Db 956121 TGAGGCAAGAGAATTCGTTGAATCCGGAGGAGAGGTTGCAGTAAGACGAGATCGCCGCC 956062
 QY 3454 ACTGGCTCCAGCTGGGCAACAGAGTGAGACCCCTGTTTCTAAAAAGAAAGAAAAA 3513
 Db 956061 ACTGCACCTCAGCTGGGCAACAGAGTGAATTCATATCCATCTCATAAAAAA 956002
 QY 3514 GGG 3516
 Db 956001 AGG 955999

RESULT 11
 PCT-US02-25201A-1/c
 ; Sequence 1, Application PC/TUS0225201A
 ; GENERAL INFORMATION:
 ; APPLICANT: APPLERA CORPORATION et al.
 ; TITLE OF INVENTION: Estrogen Receptor Alpha Variants And
 ; FILE REFERENCE: Methods Of Detection Thereof
 ; CURRENT APPLICATION NUMBER: PCT/US02/25201A
 ; PRIOR FILING DATE: 2002-08-07
 ; PRIOR APPLICATION NUMBER: 09/933,267
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 465237
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 PCT-US02-25201A-1

Query Match 7.3%; Score 317.4; DB 1; Length 465237;
 Best Local Similarity 93.5%; Pred. No. 5.1e-31;
 Matches 330; Conservative 1; Mismatches 22; Indels 0; Gaps 0;
 QY 2387 TCTTTTCATCTAAATAATAAACCATAGCGCGGGTGGCTCAGCGCTGTATCCCA 2446
 Db 47735 TATTTCTCTAAGGACAAAGAGCCAGTGGCGGGGCTCAGCGCTGTATCCCA 47676
 QY 2447 GCACCTTTGGGAGCGCGGCTGGATCATGAGTTCAGGAGATCGAGACCATCTGGCT 2506
 Db 47675 GCACCTTTGGGAGCGCGGCTGGATCATGAGTTCAGGAGATCGAGACCATCTGGCT 47616
 QY 2507 AACAGGTGAACCCCGCTCTACTTAAATAACAAAAATAGCGCGGCTGGCGGG 2566
 Db 47615 AACAGGTGAACCCCGCTCTACTTAAATAACAAAAATAGCGCGGCTGGCGGG 47556
 QY 2567 CGCCTGTAGTCCAGCTTCTCGGAGGCTGAGGCAGGAGATGCGCTGAACCCGGGAGGC 2626
 Db 47555 CGCCTGTAGTCCAGCTTCTCGGAGGCTGAGGCAGGAGATGCGCTGAACCCGGGAGGC 47496
 QY 2627 GGAGCTTCAGTGAGCCAAAGATTGGCCACTGCGAGTCCGAGTCCGCGCTGGGCGACAGA 2686
 Db 47495 GGAGCTTCAGTGAGCCAAAGATTGGCCACTGCGAGTCCGAGTCCGCGCTGGGCGACAGA 47436
 QY 2687 GCGAGACTCTCTTCAAAAAA 2739
 Db 47435 GCGAGACTCCGCTCTCAAAAAA 47383

RESULT 12
 us-60-427-045-313/c
 ; Sequence 313, Application US/60427045
 ; GENERAL INFORMATION:
 ; APPLICANT: The Board of Trustees of the University of Arkansas

; APPLICANT: O'Brien, Timothy
 ; APPLICANT: Beard, John
 ; APPLICANT: Underwood, Lowell
 ; TITLE OF INVENTION: CA125 Gene and its Use for Diagnostic and Therapeutic
 ; TITLE OF INVENTION: Interventions
 ; FILE REFERENCE: 022438.44514
 ; CURRENT APPLICATION NUMBER: US/60/427,045
 ; CURRENT FILING DATE: 2002-11-15
 ; NUMBER OF SEQ ID NOS: 314
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 313
 ; LENGTH: 14277
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)..(66)
 ; OTHER INFORMATION: Exon C1
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1802)..(1947)
 ; OTHER INFORMATION: Exon C2
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (4198)..(4350)
 ; OTHER INFORMATION: Exon C3
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (4679)..(4747)
 ; OTHER INFORMATION: Exon C4
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (6811)..(6978)
 ; OTHER INFORMATION: Exon C5
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (11232)..(11270)
 ; OTHER INFORMATION: Exon C6
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (11594)..(11677)
 ; OTHER INFORMATION: Exon C7
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (13375)..(13500)
 ; OTHER INFORMATION: n is a, c, g, or t
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (14095)..(14187)
 ; OTHER INFORMATION: Exon C8
 ; US-60-427-045-313
 Query Match 7.3%; Score 315.6; DB 7; Length 14277;
 Best Local Similarity 59.7%; Pred. No. 2.4e-30;
 Matches 668; Conservative 1; Mismatches 415; Indels 34; Gaps 7;
 QY 2400 AAAAAAACCATAGCGCGGCGGCTCAGCGCTGTATCCAGACACTTTGGGAGG 2459
 Db 8758 AAAAAAATTTGGGCTGGGTGTGGCTCAGCGCTGTATCCAGACACTTTGGGAGG 8699
 QY 2460 CCAGGCGGCTGATCATGAGTTCAGGAGATCGAGACCATCTGGCTAACAAAGGTGAAC 2519
 Db 8698 CCCAGGTGGCAGATCAGGAGTTCAGGAGATCGAGACCATCTGTGTCAACATGGTGAAC 8639
 QY 2520 CCGCTCTCTACTTAAAAATAC--AAAAATTTAGCCGGCGGCTGGCGGCGCTGTAGTC 2577
 Db 8638 CCTGTTTCTACTAAAAATACAAAAAATTTAGCCAGACATGGTGGCAGCGCTGTAGTC 8579
 QY 2578 CCAGCTTCTCGGAGGCTGAGGAGATGGCGTGAACCCGGGAGCGGAGCTTGCAG 2637
 Db 8578 CCAGCTTCTCGGAGGCTGAGGAGATGGCGTGAACCCGGGAGCGGAGCTTGCAG 8519
 QY 2638 TGAGCCAAGATTGGCGCACTGCAGTCCGCGCTGGCGGAGAGAGAGAGCTCTG 2697

```

Db 8518 TGGGGGAGATCGTGCCACT-----GCACCTCCAGCCTGGGTGACAGCGAGACTCTG 8466
Qy 2698 TTTCAAAAAAAGAAAAAAGAAAAAACCATAATGAGGAAACGCATCTTTACACTTA 2757
Db 8465 TCTCAAAAAAAGAAAAAAGAAAAAACAATAAATAAACA-AAAACAACAACAACAATAA 8407
Qy 2758 GGGTTTGAGTTCTGTATCTATAAAAAAGGGTTTGATTAAGTGTATCCTGGCACTTATA 2817
Db 8406 CAACAACAACAACAACAATAATTTGGTAACAAGAAAAAAGAAAAAAGCTGTGAC--AGG 8349
Qy 2818 AATATGTTAGGGCTTAATATTTATCATGATCGAGGATAGTTTTCATTTCTTAGTCGCTCCT 2877
Db 8348 CATGTGAAGAGCTGTGTGATTCTGAAGGATCAACAGTGTCTCTCTTGCACA----- 8299
Qy 2878 TAGTCACTCTTCTATACAACTCTGAGACCAATTTTACAAATTTAGAAAAGACAATAACTG 2937
Db 8298 ---ACAATAATTCGGAATCAAGGACAGTGCACACAGGATTTGAAGAGTCCCGTTAAAAC 8242
Qy 2938 GTTGGGTTACTTGATAGTATAAACCAGAAAAATAATTTTAGAAGGAATTAAGTTTGA 2997
Db 8241 GTTATTCGATTTGACTTGAAAGAACCTTGTATGAAATCTAGTTATTTCCCTCTCAAAAT 8182
Qy 2998 AACCACATGTTAACAAATTTCTACAAAGTGGATTTGCCTGTGATTAAGATGCTGTAAA 3057
Db 8181 ATGACTTGAGTGTATTATGTGTCAGGAGCAATATGGGGTCTTTGGAAATATAGTGGTGA 8122
Qy 3058 CATTTGGGCGAGTAGTATTAATTTGNAAAATGTTTATAGCCCAATATATAATTTTATT 3117
Db 8121 AAGAGCAAAACAGTTACTCTCTGTTATTTCTGATACATGGTACAAATGTATATACTATT 8062
Qy 3118 AATATACAGTTTCATCAGCTATTAGTATTTTCAATTAAGTCTAAGTATGACATCAGTGGTT 3177
Db 8061 CTTTGAATGGTAAATTTGGATATGCAAGTTGAGCCACCTTGAAGCCATCTCTCGAGTT 8002
Qy 3178 AGCAACA---COACTGTTTTATGCACTGCTAAGAAAGAAATAAAGGGCTGTGTGCAGTGG 3234
Db 8001 GGAATAATAGTCTCATCTACCACATTTGTTGAGGATGAATACATACAACTTTGAAACTGG 7942
Qy 3235 CTCACACCTGTGGAGCGCAAGCAGGACATCCTTGAGCGCAGGAGTCAAGACCAAC 3294
Db 7941 AAGCCCTCATAGCTGCTGAGCGAGGAGGATCATTGTGGCCAGGAATTAGAGACCACT 7882
Qy 3295 CTGGTCAACATTTGAAGACCTCTCTCTACAAAAAAGTTAAAAAATTAGCTGGGT 3354
Db 7881 CTGGCAACATAGTGAACCTCTCTCTAC-----AAAAAATTTAGAAATTAGCAAGGC 7828
Qy 3355 CGGGTGGCACATCCCTGTAGTTCCAGCTACTCTGGAGGCTAAGTGGGAGGATTTGCTAGA 3414
Db 7827 GTGGTGACAGTGCCTGTAGTTCCAGGTACTCAGGAGGCTGAGTGGGAGGATCGCTTGA 7768
Qy 3415 GCCACGCTGTTGAAGCTGCAATGAGCTGTGACCAACACCACTGCGCTCCAGCGTGGGCAA 3474
Db 7767 GCCAGAGTTCAAGACTCGAGTGAAGTATGTTGACCACTGTACTCCATTTCTGGTTAA 7708
Qy 3475 CAGAGTGAGACCTGTTTCTTAAAAGAAAGAAAAA 3512
Db 7707 CAGAGCGAGACCTGCTCTCTTAAAAGAAAAAACAACA 7670
```

```

RESULT 13
US-09-949-002-659
; Sequence 659, Application US/09949002
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; PRIOR FILING DATE: 2000-01-28
; PRIOR FILING DATE: 2000-09-08
```

```

; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 659
; LENGTH: 267482
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(267482)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-002-659

Query Match 7.2%; Score 313.8; DB 5; Length 267482;
Best Local Similarity 94.5%; Pred. No. 1.6e-30;
Matches 324; Conservative 1; Mismatches 18; Indels 0; Gaps 0;

Qy 2401 AAATAAAACCATAGGCGGGCGGGTGGCTCAGCCCTGTAATCCACGACACTTTGGGAGGC 2460
Db 258107 AAGAAAAGCAGGAGCGGGCGGGTGGCTCAGCCCTGTAATCCACGACACTTTGGGAGGC 258166
Qy 2461 CGAGCGGGTGGATCATGAGTCAAGAGATCGAGACCATCTGCTAAACAAGGTGAAACC 2520
Db 258167 CGAGCGGGTGGATCATGAGTCAAGAGATCGAGACCATCTGCTAAACAAGGTGAAACC 258226
Qy 2521 CCGTCTCTACTAAAAATACAAAAAATTAGCGGGCGGGTGGCGGCGCTGTAGTCCCA 2580
Db 258227 CCGTCTCTACTAAAAATACAAAAAATTAGCGGGCGGGTGGCGGCGCTGTAGTCCCA 258286
Qy 2581 GCTTCTCGGAGGCTGAGGAGGAGATGCGTGAACCCGAGGCGGAGCTTGCAGTGA 2640
Db 258287 GCTACTCGGAGGCTGAGGAGGAGATGCGTGAACCCGAGGCGGAGCTTGCAGTGA 258346
Qy 2641 GCCAAGATTGGCCACCTGCACTCGGCTCGGCTCGGCTGGGCGACAGCGAGCTCTGTTT 2700
Db 258347 GCCAGATTGGCCACCTGCACTCGGCTCGGCTCGGCTGGGCGACAGCGAGCTCTGTTT 258406
Qy 2701 CAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 2743
Db 258407 CAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 258449

RESULT 14
US-09-949-002-783
; Sequence 783, Application US/09949002
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; PRIOR FILING DATE: 2000-01-28
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 783
; LENGTH: 267505
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(267505)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-002-783

Query Match 7.2%; Score 313.8; DB 5; Length 267505;
Best Local Similarity 94.5%; Pred. No. 1.6e-30;
Matches 324; Conservative 1; Mismatches 18; Indels 0; Gaps 0;

Qy 2401 AAATAAAACCATAGGCGGGCGGGTGGCTCAGCCCTGTAATCCACGACACTTTGGGAGGC 2460
Db 258129 AAGAAAAGCAGGAGCGGGCGGGTGGCTCAGCCCTGTAATCCACGACACTTTGGGAGGC 258188
```

QY 2461 CGAGCGGTGATCATGAGTCAAGATCGAGACCATCTGGCTAAACAGGTGAAC 2520
 |||||
 Db 258189 CGAGCGGTGGATCATGAGTCAAGATCGAGACCATCTGGCTAAACAGGTGAAC 258248
 |||||
 QY 2521 CCGTCTCTACTATAAAATACAAAAATAGCGGCGCGGTGGCGGCGCCCTGTAGTCCCA 2580
 |||||
 Db 258249 CCGTCTCTACTATAAAATACAAAAATAGCGGCGCGGTGGCGGCGCCCTGTAGTCCCA 258308
 |||||
 QY 2581 GCTTCTCGGAGGCTGAGGAGGAGATGGCGTGAACCCCGGAGGCGGAGCTTGCAGTGA 2640
 |||||
 Db 258309 GCTACTCGGAGGCTGAGGAGGAGATGGCGTGAACCCCGGAGGCGGAGCTTGCAGTGA 258368
 |||||
 QY 2641 GCCAAGATTGGCCACTCGACTCGCGAGTCCGCGCTGGCGGACAGAGAGACTCTCTYT 2700
 |||||
 Db 258369 GCCGAGATTGGCCACTCGACTCGCGAGTCCGCGCTGGCGGACAGAGAGACTCCCTCT 258428
 |||||
 QY 2701 CAAAAAATAAAAAAAAAAAAAACCATAAATGAGGAACG 2743
 |||||
 Db 258429 CAAAAAATAAAAAAAAAAAAAAGAGCAAGG 258471
 |||||
 RESULT 15
 US-09-948-124-60/c
 ; Sequence 60, Application US/09948124
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig
 ; TITLE OF INVENTION: COLLECTION OF SINGLE NUCLEOTIDE POLYMORPHISM (SNPs) LOCATED ON CH
 ; FILE OF INVENTION: Y, METHODS OF DETECTION, AND USES THEREOF
 ; FILE REFERENCE: CL001301
 ; CURRENT APPLICATION NUMBER: US/09/948.124
 ; CURRENT FILING DATE: 2002-12-23
 ; NUMBER OF SEQ ID NOS: 183
 ; SEQ ID NO 60
 ; LENGTH: 859705
 ; TYPE: DNA
 ; ORGANISM: HUMAN
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(859705)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-948-124-60
 Query Match 7.2%; Score 312.6; DB 5; Length 859705;
 Best Local Similarity 61.8%; Pred. No. 1.6e-30;
 Matches 705; Conservative 1; Mismatches 397; Indels 37; Gaps 10;
 QY 2392 CATCTAAAAATAAACCATAGCGCGCGGTGCTCAGCGCTGTAAATCCAGCACT 2451
 |||||
 Db 584148 CACCGATAAAATAGCAATGCGCGCGGTGCTCAGCGCTGTAAATCCAGCACT 584089
 |||||
 QY 2452 TTGGAGCGCGAGGCGGTGATCATGAGTCAAGAGATCGAGACCATCTGGCTAACAA 2511
 |||||
 Db 584088 TTGGAGACCGAGCGCGGATCAGAGGTCAAGAGATCGAGACCATCTGGCTAACAA 584029
 |||||
 QY 2512 GTGAAACCCCTCTCTACTAAAAATACAAAAATAGCGGCGCGGTGCGGCGCCT 2571
 |||||
 Db 584028 GTGAAACCCCTCTCTACTGAAATACAAAAATAGCGGCGGTGATAGCGGCGCCT 583969
 |||||
 QY 2572 GTAGTCCCAAGCTTCTCGGAGGCTGAGGAGGAGATGCGGTGAACCCGGAGCGGAGC 2631
 |||||
 Db 583968 GTAGTCCCAAGCTTCTCGGAGGCTGAGGAGGAGATGCGGTGAACCCGGAGCGGAGC 583909
 |||||
 QY 2632 TTGCAGTGAAGCAATGCGGCCTGAGTCCGAGTCCGCGCTGGCGGACAGAGCGAG 2691
 |||||
 Db 583908 TTGCAGTGAAGCAATGCGGCCTGAGTCCGAGTCCGCGCTGGCGGACAGAGCGAG 583856
 |||||
 QY 2692 ACTCTCTCAAAAAAATAAAAAAAAAAAAAACCATAAATAGGAACCGCATCTTTA 2751
 |||||
 Db 583855 ACTCTCTCAAAAAAATAAAAAAAAAAAGAGCAATGAGAAAGACTCCCTATTCAACTAAC 583796
 |||||
 QY 2752 CACTTAGGTTTTCAGTTTCTCTATATAAAAAAGGTTTGGATTAAAGTATCCCTGGCA 2811
 |||||

Db 583795 AGTGTGGTGGTAATAACTTGTCTAGCCACACAGAGAAATGAA--ACTGGATCCTTAOCT 583738
 QY 2812 CTTATAAAATGTTAGGGCTTAATAATTATTATCATAGATCGAGGATAGTTTCATTCTTAGTCG 2871
 |||||
 Db 583737 TTCACACATATAAAATTAACCTAAGATGGACTAAATATTTAAATGT--ACGATGTCNA 583680
 |||||
 QY 2872 CTTCTTCTAGTCACTCTTCTATATACCAATCTGAGACCAATTTTACAAATTTAGAAAAGACAAA 2931
 |||||
 Db 583679 ACTATAAAAAAGAAAAACCTAGAAAAATACCTTCTCAACGCTGGTTTTCGAAAAACAATTT 583620
 |||||
 QY 2932 TAACTGGTTGGGTACTTCTGATAGTATAATAACCAAGAAAAATAAT--TTTAGAAGAAAT 2989
 |||||
 Db 583619 TTAACATATGTCCTTAATAATCAACTGCAACAAAAATGAAAATGAAAGTGGGACCCCAAT 583560
 |||||
 QY 2990 AAGTTTGAACCAACATGTTTAAACAAATCTACCAAAAGTGGGATTTGCTGTGATTTAAAGAT 3049
 |||||
 Db 583559 CAGGTATGAGCTTTCAGCAGACGCAAAATTAACAAACAAAAATATCAACAGAGTAAACAG 583500
 |||||
 QY 3050 GCTGTAAACATTTGGGCCAGTAGTTATATAATTTGAAAAATGTTTATAGCAATATATAATT 3109
 |||||
 Db 583499 ACAAGCTACAAAATGGGATAAATTAATG--GTGAACATATGATCTGACCAAGTCTAATA 583442
 |||||
 QY 3110 TTTTATTAAATATACAGTTTTCATCAGTCTATTAGTATTTCATTAA-GTCTAAGATGCCA 3168
 |||||
 Db 583441 TCAGAAATCTACAAAGAACTTAAACAAATCAACAAAGAAAAACAAATGACCGCACTAAAA 583382
 |||||
 QY 3169 TCAGTGGTTAGCAAAACACCACTGTTTATGCACTGCTAAGA-AAGAAATAAAGGGCTGTGT 3227
 |||||
 Db 583381 AATGGACAAGGTTATGAANAAGGTATGCTACTAAAAGAGGACATTAACAGGCGCGGC 583322
 |||||
 QY 3228 GCAGTGGCTCACACCTG-----TGGGACGCCAAGGCGAGGAGCATCACTTGA 3273
 |||||
 Db 583321 GCGGTGGCTCACGCTGTAAATCCAGCACTTTGGGAGCGCGAGCGGGGATCAC--GA 583264
 |||||
 QY 3274 GGCAGAGTTCAAGACCAACCTGGTCAACATTTGAAGACCTGTCTCTACAAAAAATA 3333
 |||||
 Db 583263 GGTCAAGAGATCGAGACCATCTTAGTAAACAGAGTGAACAGAGTGAACCCGCTCTCTACTAAAAATA 583204
 |||||
 QY 3334 AAAATTAAAAATTAGCTGGTGGCGGTGGCACATGCTGTAGTTCCAGCTACTCTGGAGGC 3393
 |||||
 Db 583203 CA----AAAAATTAGCGGCGGTGGTGGCGGCGCTGTAGTCCAGCTACGCGGGAGGC 583148
 |||||
 QY 3394 TAAGGTGGGAGGATTTGTAGAGCCACGGTGTGGAAGCTGCAATGAGCTGTGACCAACC 3453
 |||||
 Db 583147 TGAGGCAAGGAGATGGCGTGAACCCGGAGGCGGAGCTTGCAGTGAGCCGAGATCGCGC 583088
 |||||
 QY 3454 ACTGCGCTCCAGCGTGGCAACAGAGTGAACCCCTGTTTCTAAAGAAAGAAAAAATA 3513
 |||||
 Db 583087 ACTGCACTCCAGCCTGGGCGACAGAGCGGAGACTCCGCTCTCAAAAAAAGAGGACATA 583028
 |||||

Search completed: February 28, 2003, 14:54:38
 Job time : 21575 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run On: February 28, 2003, 00:23:22 ; Search time 5553 Seconds
(without alignments)
12666.486 Million cell updates/sec

Title: US-09-730-559B-7
Perfect score: 4343
Sequence: 1 atggacgcccccaagcagg.....attcacagcagtcctcttct 4343

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- EST: *
- 1: em_estba:*
 - 2: em_esthum:*
 - 3: em_estin:*
 - 4: em_estmu:*
 - 5: em_estcov:*
 - 6: em_estpl:*
 - 7: em_estro:*
 - 8: em_htc:*
 - 9: gb_est1:*
 - 10: gb_est2:*
 - 11: gb_htc:*
 - 12: gb_est3:*
 - 13: gb_est4:*
 - 14: gb_est5:*
 - 15: em_estfun:*
 - 16: em_estom:*
 - 17: gb_gss:*
 - 18: em_gss_hum:*
 - 19: em_gss_inv:*
 - 20: em_gss_pln:*
 - 21: em_gss_vrt:*
 - 22: em_gss_fun:*
 - 23: em_gss_mam:*
 - 24: em_gss_mus:*
 - 25: em_gss_other:*
 - 26: em_gss_pro:*
 - 27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	794.6	18.3	875	14	BQ433500
2	717.4	16.5	954	14	BM802385
3	648.6	14.9	1068	12	BG260015
4	626.8	14.4	938	12	BG387889
5	616.6	14.2	1076	14	BQ224816
6	578	13.3	664	10	AV759698

7	510	11.7	754	12	BF669398
8	494.2	11.4	581	14	N63786
9	487.4	11.2	509	9	AA224499
10	485.6	11.2	567	9	AA104450
11	482.8	11.1	498	9	AA313478
12	463.8	10.7	467	9	AL121430
13	444	10.2	865	12	BF575050
14	426.4	9.8	967	12	BG258171
15	410.2	9.4	489	14	N94143
16	408.4	9.4	421	14	BQ333861
17	407.2	9.4	681	10	BE376722
18	406.6	9.4	826	9	AJ456404
19	390	9.0	398	10	AM470799
20	379.4	8.7	419	10	AV761302
21	376.4	8.7	436	14	H68318
22	372.6	8.6	773	9	AJ455338
23	340.4	7.8	371	14	BQ333859
24	326	7.5	626	9	AJ454445
25	316.2	7.3	698	9	AJ452920
26	314.2	7.2	1099	14	BM799574
27	312	7.2	507	13	BI494159
28	311.8	7.2	475	17	AQ474970
29	310	7.1	452	13	BI494158
30	309	7.1	501	17	AQ799583
31	308.6	7.1	669	17	AQ108614
32	308.4	7.1	549	17	B55844
33	308	7.1	661	17	AQ415932
34	307.8	7.1	583	13	BI438426
35	307.2	7.1	606	12	BF684828
36	306.2	7.1	578	17	AQ506712
37	306	7.0	392	17	AQ285940
38	305.8	7.0	581	9	AL702636
39	305.6	7.0	353	13	BI494884
40	305.4	7.0	679	14	BQ025061
41	304.8	7.0	636	9	AL038785
42	304.8	7.0	835	17	AQ48746
43	304.6	7.0	352	13	BI494883
44	304.6	7.0	442	17	AQ262563
45	304.6	7.0	724	17	B49481

ALIGNMENTS

RESULT 1
BQ433500
LOCUS
DEFINITION
5', mRNA sequence.
AGENCOURT_7827311 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:6014314
875 bp mRNA linear EST 24-MAY-2002
ACCESSION
BQ433500
VERSION
BQ433500.1
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 875)
NITH-MGC <http://mgc.nci.nih.gov/>.
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE
Unpublished (1999)
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLAM3208
row: j
column: 11
High quality sequence stop: 652.
Location/Qualifiers
1. .875
source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6014314"
/clone_lib="NIH_MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/site_1: NotI;
/site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT      251 a  105 c  197 g   242 t
ORIGIN

Query Match      18.3%; Score 794.6; DB 14; Length 875;
Best Local Similarity 99.2%; Pred. No. 7.8e-89;
Matches 820; Conservative 0; Mismatches 4; Indels 3; Gaps 2;

QY  1  ATGACGCCCCAAAGCAGGATAGCGCTTTGAGTACCTTATTGGAACATTAAATGACAGT 60
Db   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  61 TCACATAGAAGTTCTTCGATGATCTATAACTTGGCACCAGATGATGTTCTCGCCTTAC 120
Db   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  110 TCACATAGAAGTTCTTCGATGATCTATAACTTGGCACCAGATGATGTTCTCGCCTTAC 169
Db   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  121 TCAATACGGGTCTTGTGGAGCTGCTGTACGAAATTTGTATGGCTTTTAAATGAAGAAG 180
Db   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  170 TCAATACGGGTCTTGTGGAGCTGCTGTACGAAATTTGTATGGCTTTTAAATGAAGAAG 229
Db   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  181 GAAGATGTTATGAACATTTTAGACTGGAACCAACCAAGCAATGTTGAAGTGCCTTTT 240
Db   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  230 GAAGATGTTATGAACATTTTAGACTGGAACCAACCAAGCAATGTTGAAGTGCCTTTT 289
Db   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  241 TTCCCTGCCGTGTTCTTCAAGATTTTACTGGAATACCAGCAATGTTGATTTTGGT 300
Db   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  290 TTCCCTGCCGTGTTCTTCAAGATTTTACTGGAATACCAGCAATGTTGATTTTGGT 349
Db   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  301 GCTATGAGGAGCAGTGAACCTCTTGGAGTGATCCTGAGAAAGTCCATCCTGTTGT 360
Db   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  350 GCTATGAGGAGCAGTGAACCTCTTGGAGTGATCCTGAGAAAGTCCATCCTGTTGT 409
Db   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  361 CCAGACAGATCTACAGTTGACCATTTCTTACAATTTGACTTCAGTAAATGTCGAATACAG 420
Db   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  410 CCAGACAGATCTACAGTTGACCATTTCTTACAATTTGACTTCAGTAAATGTCGAATACAG 469
Db   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  421 AATGCACCAATCCTGGAGTGTGACCTGCAGAAAGCAGGAAAGCTCTCTCCACTTAA 480
Db   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  470 AATGCACCAATCCTGGAGTGTGACCTGCAGAAAGCAGGAAAGCTCTCTCCACTTAA 529
Db   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  481 GTGAGCCTTAAGAAGCTTCCCTGCAGAGGCCAGACTACCTGCCGAGGATCTTGTGATTC 540
Db   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  530 GTGAGCCTTAAGAAGCTTCCCTGCAGAGGCCAGACTACCTGCCGAGGATCTTGTGATTC 589
Db   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  541 GGAAACTAGGCCGAACCTCAGAAATTTTCTCCAGATTCAGAAATACCCATCTCG 600
Db   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  590 GGAAACTAGGCCGAACCTCAGAAATTTTCTCCAGATTCAGAAATACCCATCTCG 649
Db   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  601 TGTCTCTTTTCATTTGCAACACAGTGCCTGAACTGAAACAGTGTATAAAATCAAGAAGTA 660
Db   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  650 TGTCTCTTTTCATTTGCAACACAGTGCCTGAACTGAAACAGTGTATAAAATCAAGAAGTA 709
Db   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  661 GAATTCGGCAGAAATCGAGAGGCTTCAGTTTTTAACTGAGTTCAAGAGTTTAAAG 720
Db   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  710 GAATTCGGCAGAAATCGAGAGGCTTCAGTTTTTAACTGAGTTCAAGAGTTTAAAG 769
Db   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  721 AATGTGCAGTATCCTCTCGAACTGGAATGGCTCATCAATAAATTAAGTAAATTTG 780
Db   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  770 AATGTGCAGTATCCTCTCGAACTGGAATGGCTCATCAATAAATTAAGTAAATTTG 828
Db   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  781 TCAAGAGT--GGTTTTTGAAGAAAGACCTCTCTCTCCAGACAGT 825
Db   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  829 TCAAGAGTGGTTTTTGAAGAAAGACCTCTCTCTCCAGAAAGT 875
Db   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

RESULT 2

BM802385

LOCUS

DEFINITION

AGENCY

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

Location/Qualifiers

1..954

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5559954"

/clone_lib="NIH_MGC_88"

/tissue_type="duodenal adenocarcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: small intestine; Vector: pCMV-SPORT6;"

Site_1: NotI; Site_2: SalI; Cloned unidirectionally;"

oligo-dT primed. Average insert size 1.767 kb. Library

enriched for full-length clones and constructed by Life

Technologies. Note: this is a NIH_MGC Library."

308 a 189 c 218 g 238 t 1 others

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches 863; Conservative 1; Mismatches 67; Indels 26; Gaps 8;

QY 2399 AAAATAAAACCATAGCGCGCGCGGTGGCTCACGCTGTAAATCCAGCAGTCTTGGGAG 2458

Db 1 AAAATAAAACCATAGCGCGCGCGGTGGCTCACGCTGTAAATCCAGCAGTCTTGGGAG 60

QY 2459 GCCGAGCGGGTGGATCATGAGGTTCAGAGATCGAGACCATCTCTGGCTTACAGAGTGA 2518

Db 61 GCCGAGCGGGTGGATCATGAGGTTCAGAGATCGAGACCATCTCTGGCTTACAGAGTGA 120

QY 2519 CCCGCTCTCTACTTAAATACAAAAATAGCCGGCGCGGTGGCTGGCTTACAGAGTGA 2578

Db 121 CCCGCTCTCTACTTAAATACAAAAATAGCCGGCGCGGTGGCTGGCTTACAGAGTGA 180

QY 2579 CAGCTTCTCGGAGGCTGAGCAGGAGAAATGGCTGAACCCGGGAGCGGAGCTTTCAGT 2638

Db 181 CAGCTTCTCGGAGGCTGAGCAGGAGAAATGGCTGAACCCGGGAGCGGAGCTTTCAGT 240

QY 2639 GAGCAAGATTCGGCAGTCCGAGTCCGCGCTCGGCGTGGCGGACAGAGGAGACTCTGT 2698

Db 241 GAGCGGATTCGCCACTTCAGTCCGAGTCCGCGCTCGGCGTGGCGGACAGAGGAGACTCTGT 300

QY 2699 YTCAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 2758

Db 301 CTC--AAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 356

```
QY 2759 GGTGTTGAGTTCTGTATCTATAAAAAAGGTTTGGATTAAAGTATCCCTGGCAGCTATAA 2818
Db 357 GGTGTTGAGTTCTGTATCTATAAAAAAGGTTTGGATTAAAGTATCCCTGGCAGCTATAA 416
QY 2819 AATGTTAGGCTTAATATTATTCATAGATCGAGGATAGTTTCATTCTTAGTCGCCTCCTT 2878
Db 417 AATGTTAGGCTTAATATTATTCATAGATCGAGGATAGTTTCATTCTTAGTCGCCTCCTT 476
QY 2879 AGTCACTCTCTCTATACCAATCTGAGACCAATTTTCAATTTTGAAGAACAAATAACATGG 2938
Db 477 AGTCACTCTCTCTATACCAATCTGAGACCAATTTTCAATTTTGAAGAACAAATAACATGG 536
QY 2939 TTGGTTACTTGTATAGTATAAACCAGAAAAATATTTTGAAGGAATTAAGTTTGA 2998
Db 537 TTGGTTACTTGTATAGTATAAACCAGAAAAATATTTTGAAGGAATTAAGTTTGA 596
QY 2999 ACCACATGTTACAAATCTACCAAGTGGGATTTTGCCTGTGATTTAAAGATGCTGTAAC 3058
Db 597 ACCACATGTTACAAATCTACCAAGTGGGATTTTGCCTGTGATTTAAAGATGCTGTAAC 656
QY 3059 ATTTGGGCCAGTAGTTATATTTGAAAAATGTTTATAGCCATATATATTTTATTTA 3118
Db 657 ATTTGGGCCAGTAGTTATATTTGAAAAATGTTTATAGCCATATATATTTTATTTA 716
QY 3119 AATATACAG-TTTCATCAGT-CTATTAGTATTTCATTAGTCTAAGATGCCATCAGTGGT 3176
Db 717 AATATAAAGTTTTCATCAGTCTATTAGTATTTCATTAGTCTAAGATGCCATCAGTGG 776
QY 3177 TAGCAACACAC- ---TGTTTATGCACTGCTAAGAAAGATAAAGGCTGTGTGCACT 3232
Db 777 GTTAGCACACACACCTGGTGTGTTTATGCACTGCTAAGAAAGATAAAGGCTGTGTG 836
QY 3233 G-----GCTCACACCTGTGGAGCGCAAGCGAGGAGCATCA- ---CTTAGAGGCA 3278
Db 837 GCCAATGGCATCCACCCCTGTGGAGCGCCAGGAGGTCAGGAGGCCATCACTTTAAGGGCA 896
QY 3279 GAAGTTCAAGACCAAC- ---TGCTACACATGTTAAGACCCCTGTCTCTACAAAAA 3333
Db 897 GAAGATCCAAACACCCCGGGGTCAACATTTGAAAAACCTGTCTCTAACAACAAAA 953

RESULT 3
BG260015
LOCUS
DEFINITION
602371792F1 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:4479791 5',
mRNA sequence.
ACCESSION
BG260015
VERSION
BG260015.1 GI:12769831
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1068)
NIH-MGC http://mgc.nci.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10312 row: g column: 24
High quality sequence start: 2
High quality sequence stop: 655.
Location/Qualifiers
1..1068
/organism="Homo sapiens"
/db_xref="taxon:9606"
```

```
BASE COUNT 366 a 181 c 254 g 267 t
ORIGIN
Query Match 14.9%; Score 648.6; DB 12; Length 1068;
Best Local Similarity 85.9%; Pred. No. 5.8e-71;
Matches 779; Conservative 0; Mismatches 119; Indels 9; Gaps 5;

QY 208 AAAACCAACAAACAAATGTTGAAGTGCCTTTTTCCTCC- ---TGCCCGTGTCTTCTTCA 263
Db 1 AAAACCAACAAACAAATGTTGAAGTGCCTTTTTCCTCCCTGGCGCGGTGTCTTCTTCA 60

QY 264 AGATTTTACTGGAATACCAAGCAATGGTGGATTTTGTCTTATGAGGAGCGAGTGAAC 323
Db 61 AGATTTTACTGGAATACCAAGCAATGGTGGATTTTGTCTTATGAGGAGCGAGTGAAC 120

QY 324 TCTTGGAGGTGATCCTCGAGAAAGTCCATCCTGCTGTGTCGACAGATCTTACAGTTGACCA 383
Db 121 TCTGGAGGTGATCCTCGAGAAAGTCCATCCTGCTGTGTCGACAGATCTTACAGTTGACCA 180

QY 384 -TCTTTTACAAATGACTTCAGTAAATGTGCAATACAGAAATGACCAAAATCCTGGAGGTG 442
Db 181 GTTCTTTACAAATGACTTCAGTAAATGTGCAATACAGAAATGACCAAAATCCTGGAGGTG 240

QY 443 GTGACCTGCAGAAAGCAGGAAGCTCTCCACCTTAAGTGCAGCCTTAAGCAAGCTCCCT 502
Db 241 GTGACCTGCAGAAAGCAGGAAGCTCTCCACCTTAAGTGCAGCCTTAAGCAAGCTCCCT 300

QY 503 GCAGAGGCCAGCTACCTGCGGAGGATCTGTGATTTCTGGAGAACTAGGCGGAACTCAG 562
Db 301 GCAGAGGCCAGCTACCTGCGGAGGATCTGTGATTTCTGGAGAACTAGGCGGAACTCAG 360

QY 563 GAACATTTTCTGCGAGATTGAGAAATACCCATPCCTGTGTCTTTTTCATTGTCAGCAG 622
Db 361 GAACATTTTCTGCGAGATTGAGAAATACCCATPCCTGTGTCTTTTTCATTGTCAGCAG 420

QY 623 TGCTGACCTGGAACAGTGTGTTAAATAATCAAGAAGTAGAATTCGGCAGAAATCGAGAGA 682
Db 421 TGCTGACCTGGAACAGTGTGTTAAATAATCAAGAAGTAGAATTCGGCAGAAATCGAGAGA 480

QY 683 GGCTTCAGTTTTTTAAGTGGAGTTCAAGAGTTTTTAAAGAAATGTCGAGTATCCCTCCTG 742
Db 481 GGCTTCAGTTTTTTAAGTGGAGTTCAAGAGTTTTTAAAGAAATGTCGAGTATCCCTCCTG 540

QY 743 GAATCGAATGGCTCATCAATAAATAGTAATTTGTCAGAGTGGTTTTTGAAGAA- 801
Db 541 GAATCGAATGGCTCATCAATAAATAGTAATTTGTCAGAGTGGTTTTTGAAGAA- 600

QY 802 AAACACCTCTCTTCCAGCAGTGTAGTGGCAGACATTCACACATAAGCATGCTCAAT 861
Db 601 AAACACCTCTCTTCCAGCAGTGTAGTGGCAGACATTCACACATAAGCATGCTCAAT 660

QY 862 GGTTTAGGATTCCTGGGTTGGGGTTGGAGGATTTGAAACAGAAAGCATGTTAT-GCTTGG 920
Db 661 GGTTTAGGATTCCTGGGTTGGGGTTGGAGGATTTGAAACAGAAAGCATGTTATAGCAGGA 720

QY 921 TCTGCCAGTTCTTCTTACTTTACAGAGGTTTGGATGTGAGTTAACTGGTGCATCAAA 980
Db 721 CTGCGCAATTTCTTTAAIGTACACAAGGTTGTCATTTGAATTA- -TCTGCATATAA 778

QY 981 CCCTTTTCTTACATCCATAGATGTTCTTCTGTTATTTACAAAGTAAAGTTTAACTTCTGG 1040
Db 779 ACCTTTTCTTAAATCCATACAGTGTGCTGGATATAAAGGACATAGTGAATAAGGG 838

QY 1041 TAGCTCTATGACTTACTGAACATTTATTTTATAAAAAATTTGAAGAGCTCTATAGAGCAGG 1100
```

```
/clone-"IMAGE:4479791"
/clone_lib="NIH_MGC_93"
/tissue_type="transitional cell papilloma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: bladder; vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
```

```

Db      839  GGACAAAGTGAAGACGATATTTTATATTTTCTGAAAAAGATATAATGAAG 898
Qy      1101  GATTGG 1107
Db      899  AGTGAGG 905

RESULT 4
LOCUS    BG387889
DEFINITION 602412909F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4521122 5',
mRNA sequence.
VERSION  BG387889
KEYWORDS EST.
SOURCE    BG387889.1 GI:13281335
          human.
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10420 row: b column: 03
High quality sequence stop: 631.

FEATURES
    source
        1..938
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="IMAGE:4521122"
        /clone_lib="NIH_MGC_92"
        /tissue_type="embryonal carcinoma, cell line"
        /lab_host="DH10B (phage-resistant)"
        /note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;
        Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
        Average insert size 2.5 kb. Library enriched for
        full-length clones and constructed by Life Technologies.
        Note: this is a NIH_MGC Library."
BASE COUNT 300 a 250 c 195 g 193 t
ORIGIN
    Query Match 14.4%; Score 626.8; DB 12; Length 938;
    Best Local Similarity 90.9%; Pred. No. 3e-68;
    Matches 700; Conservative 0; Mismatches 67; Indels 3; Gaps 3;

Qy      1  ATGGAGCCCCAAAGCAGGATACGCTTTGAGTACCTATTGAAACATTAATGACAGT 60
        |||||||
Db      50  ATGGAGCCCCAAAGCAGGATACGCTTTGAGTACCTATTGAAACATTAATGACAGT 109
        |||||||
Qy      61  TCACATAAGAAGTCTTCATGATATCTAACTTGCACCACCAAGTATGATGCTGCTTAC 120
        |||||||
Db      110  TCACATAAGAAGTCTTCATGATATCTAACTTGCACCACCAAGTATGATGCTGCTTAC 169
        |||||||
Qy      121  TCACATACGGGTCTTGTGGAAGCTGCTGTACGAAATTTGATGGCTTTTATGAAGAAG 180
        |||||||
Db      170  TCACATACGGGTCTTGTGGAAGCTGCTGTACGAAATTTGATGGCTTTTATGAAGAAG 229
        |||||||
Qy      181  GAAGATGTTATGAACATTTTAGACTGGAAACCAACAAAGCAATGTTGAAGTGCCTTT 240
        |||||||
Db      230  GAAGATGTTATGAACATTTTAGACTGGAAACCAACAAAGCAATGTTGAAGTGCCTTT 289
        |||||||
Qy      241  TTCCCTGCCCGTCTCTTCAAGATTTTACTGGAATACCAAGCAATGGTGGATTTTCT 300
        |||||||

```

```

Db      290  TTCCCTGCCCGTCTCTTCTTCAAGATTTTACTGGAATACCAAGCAATGGTGATTTTGCT 349
Qy      301  GCTATGAGGAGGAGTGAAGAACTCTTTGGAGGTGATCCTGAGAAAGTCCATCTCGCTTGT 360
        |||||||
Db      350  GCTATGAGGAGGAGTGAAGAACTCTTTGGAGGTGATCCTGAGAAAGTCCATCTCGCTTGT 409
        |||||||
Qy      361  CCGACAGATCTTACAGTTGACCATCTTTTACAAATTTGACATTAATGTCGAATACAG 420
        |||||||
Db      410  CCGACAGATCTTACAGTTGACCATCTTTTACAAATTTGACATTAATGTCGAATACAG 469
        |||||||
Qy      421  AATGCACCAAAATCTGGAGGTGGTACCTGCAGAAAGCAGGAAAGCTCTCTCCACTTAAA 480
        |||||||
Db      470  AATGCACCAAAATCTGGAGGTGGTACCTGCAGAAAGCAGGAAAGCTCTCTCCACTTAAA 529
        |||||||
Qy      481  GTGCAGCCTTAAGAAGCTTCCCTGCAGAGGCCAGACTACCTGCCGAGGATCTTGTGATTCT 540
        |||||||
Db      530  GTGCAGACTAAGAAGCTTCCCTGCAGAGGCCAGACTACCTGCCGAGGATCTAGTGATTCT 589
        |||||||
Qy      541  GGAGAACTAGGCCGAAACTCAGGAACATTTTCTCGCAGATTGAGATACACCATCTG 600
        |||||||
Db      590  GGAGAACTAGGCCGAAACTCAGGAACA-TTCTTTCGACAGATTGAGAAATACAGCCATCTG 648
        |||||||
Qy      601  TGTCTTTTTCATTTGCAACACAGTGCCTGAACTGAAACAGT-GTTAAAAATCAAGAAGT 659
        |||||||
Db      649  TGTCTTTTACATTAGCAAAACAGTGCCTGAACTGACACAGTGGTGCACAAATCAAGACGC 708
        |||||||
Qy      660  AGAATTCGGCAGAAATCGAGAGAGGCTTCAGTTTTTAAAGTGGAGTTCAAG-AGTTTTAA 718
        |||||||
Db      709  ACACATCGGCCGACATCTGAGAAAGCCAGACACAGACAGAACTCAGCAAGTACAC 768
        |||||||
Qy      719  AGAATGTGGCAGTATCCCTCCTCGAACTGGAATGGCTCATCAATAAATAC 768
        |||||||
Db      769  GAAATGTGCGAGGCGACCCCGGAAACGACAGCCCAACACCAACGACAC 818
        |||||||

RESULT 5
LOCUS    BQ224816
DEFINITION  BQ224816 1076 bp mRNA linear EST 02-MAY-2002
5', mRNA sequence.
ACCESSION  BQ224816
VERSION    BQ224816.1 GI:20406226
KEYWORDS  EST.
SOURCE    human.
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCRD/DTP/Gazdar
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13323 row: j column: 03
High quality sequence start: 93
High quality sequence stop: 383.

FEATURES
    source
        1..1076
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="IMAGE:6058466"
        /clone_lib="NIH_MGC_68"
        /tissue_type="large cell carcinoma"
        /lab_host="DH10B (phage-resistant)"
        /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
        Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.

```

Average insert size 1.8 kb. Library constructed by Life Technologies."

BASE COUNT 297 a 234 c 263 g 282 t

Query Match 14.28; Score 616.6; DB 14; Length 1076;
Best Local Similarity 98.68; Pred. No. 4.9e-67;
Matches 622; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 393 AATTGACCTTCAGTAATGTCATACAGAAATGACCAAAATCCTGGAGGTGGTGAACCTGCA 452
Db 1 AATTGACCTTCAGTAATGTCATACAGAAATGACCAAAATCCTGGAGGTGGTGAACCTGCA 60

QY 453 GAAACGAGGAAGCTCTCCACATTAAGTGCAGCCTACAGACCTCCCTGCAGAGGCCA 512
Db 61 TAAACGAGGAAGCTCTCCACATTAAGTGCAGCCTACAGACCTCCCTGCAGAGGCCA 120

QY 513 GACTACCTGCCAGAGATCTTGATCTGGAGAACTAGGCCGAACTCAGGAACATTTTC 572
Db 121 GACTACCTGCCAGAGATCTTGATCTGGAGAACTAGGCCGAACTCAGGAACATTTTC 180

QY 573 TTCGAGATGAGAAATACACCATCTGTCTCTTTTCATTTGCAACCAAGTCCCTGAACC 632
Db 181 TTCGAGATGAGAAATACACCATCTGTCTCTTTTCATTTGCAACCAAGTCCCTGAACC 240

QY 633 TGAACAGTGTAAAAATCAAGATAGAAATTCGGCAGAAATCGAGAGAGCTTCAGTT 692
Db 241 TGAACAGTGTAAAAATCAAGATAGAAATTCGGCAGAAATCGAGAGAGCTTCAGTT 300

QY 693 TTTTAAGTGGAGTCAAGAGTGTAAAGATGTGGCAGTGTATCCCTCCCTGGAACCTGGAAT 752
Db 301 TTTTAAGTGGAGTCAAGAGTGTAAAGATGTGGCAGTGTATCCCTCCCTGGAACCTGGAAT 360

QY 753 GGCATCAATAAATCTAGAAATTTGTCAGAGTGGTTTGAAGAAAAAGACCTCCT 812
Db 361 GGCATCAATAAATCTAGAAATTTGTCAGAGTGGTTTGAAGAAAAAGACCTCCT 420

QY 813 CTTCCAGACAGTGTAGTCGCGCAGATTCACACATTAACGATGGTGAATGGTTAGGGAT 872
Db 421 CTTCCAGACAGTGTAGTCGCGCAGATTCACACATTAACGATGGTGAATGGTTAGGGAT 480

QY 873 TCTGGGTGGGGGTGGAGCATTTGAACAGAGCAGTATGCTTGGTCTCCAGTTTC 932
Db 481 TCTGGGTGGGGGTGGAGCATTTGAACAGAGCAGTATGCTTGGTCTCCAGTTTC 540

QY 933 TCTTACTTTACAGAGTGGTGGATGTGAGTTAACTGGTGCATCAAAACCCCTTTGTTAC 992
Db 541 TCTTACTTTACAGAGTGGTGGATGTGAGTTAACTGGTGCATCAAAACCCCTTTGTTAC 600

QY 993 ATCCATAGATGTGTTCTTGGTATTACAAAG 1023
Db 601 ATCCATAGATGTGTTCTTGGTATTACAAAG 631

RESULT 6 664 bp mRNA linear EST 19-OCT-2000
LOCUS AV759698 MDS Homo sapiens cdna clone MDSBBC05 5', mRNA sequence.
DEFINITION AV759698
ACCESSION AV759698
VERSION AV759698.1 GI:10917546
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 664)
AUTHORS Gu.J., Zhao.M., Huang.Q., Xu.X., Li.Y., Peng.Y., Song.H., Xiao.H.,
Li.N., Qian.B., Liu.F., Qu.J., Gao.X., Cheng.Z., Xu.Z., Zeng
L., Xu.S., Gu.W., Tu.Y., Jia.J., Fu.G., Ren.S., Zhong.M., Lu.G.,
Yang.Y., Gao.G., Zhang.Q., Chen.S., Han.Z. and Chen.Z.
TITLE Homo sapiens cdna MDS clones
JOURNAL Unpublished (2000)
COMMENT Contact: Zeguang Han

Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801923
Email: hanzg@hgc.sh.cn
This clone is available at CHGC in Shanghai.
Location/Qualifiers
1. 664
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MDSBBC05"
/clone_lib="MDS"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/note="vector: pTriplEx2; Site_1: sfliA; Site_2: sfliB"
BASE COUNT 191 a 110 c 107 g 247 t 9 others
ORIGIN

Query Match 13.38; Score 578; DB 10; Length 664;
Best Local Similarity 94.58; Pred. No. 3.6e-62;
Matches 617; Conservative 0; Mismatches 29; Indels 7; Gaps 2;

QY 1693 AACTTCAATACATACACTGAGCTTGTAAAGCATATTAAATACAGAGCTCAGCGATTTCCT 1752
Db 1 AGCTTCAATACATACACTGAGCTTGTAAAGCATATTAAATACAGAGCTCAGCGATTTCCT 60

QY 1753 AGTGAACAATAATTTGTAACCTCTTCTTCCCTAAATGCTGGCCCTTTGCTAACCTTTATTTA 1812
Db 61 AGTGAACAATAATTTGTAACCTCTTCTTCCCTAAATGCTGGCCCTTTGCTAACCTTTATTTA 120

QY 1813 ATGATTAATCCTATTTGTTAAATGAATGATCTACCTGGAATAATGTTCCACATATTAATCCA 1872
Db 121 ATGATTAATCCTATTTGTTAAATGAATGATCTACCTGGAATAATGTTCCACATATTAATCCA 180

QY 1873 ATTTGAGTCCCAATCTCAGCATTTTGGTTAGATATTGTTAGAGGCTTTCTGGATAC 1932
Db 181 ATTTGAGTCCCAATCTCAGCATTTTGGTTAGATATTGTTAGAGGCTTTCTGGATAC 240

QY 1933 TCCAGTGTAAAGAAATGATAGTCCCTCCCTCAGCATTTGGTATGATGCTTCCTCCT 1992
Db 241 TCCAGTGTAAAGAAATGATAGTCCCTCCCTCAGCATTTGGTATGATGCTTCCTCCTCCT 300

QY 1993 AATTAGAAAGAAATTTGGCATCTTTAGAGAAATATTGATTCAACGTATGATACCAAGA 2052
Db 301 AATTAGAAAGAAATTTGGCATCTTTAGAGAAATATTGATTCAACGTATGATACCAAGA 360

QY 2053 TCAAGTGTAAATTTGGAAATTCAGAGATTTCTCTAGAGAAAGAGGATCCCATATG 2112
Db 361 TCAAGTGTAAATTTGGAAATTCAGAGATTTCTCTAGAGAAAGAGGATCCCATATG 420

QY 2113 TTTTACAGAAATCAATCTTTTACTTTAGACATCTCTGAAACCTAACGCTGCTTTTAGCC 2172
Db 421 TTTTACAGAAATCAATCTTTTACTTTAGACATCTCTGAAACCTAACGCTGCTTTTAGCC 480

QY 2173 TTTCTAGCTGTTTTCTTCCACAAATATTACTGTGTGTTTTTTTGACATTTAGTTAATG 2232
Db 481 TTTCTAGCTGTTTTCTTCCACAAATATTACTGTGTGTTTTTTTGACATTTAGTTAATG 540

QY 2233 TTAAAAAATTAATCTATTATATATGTTTACATATTGATGATATGATCTCTTTT 2292
Db 541 GT---AAATTAATCTATTATATATGTTTACATATTGATGATATGATCTCTTTTNN 593

QY 2293 TCAGATCTGTTTCCATTTGTCATCTTTATAGGAATATCTCTGATGTTGTTTTT 2345
Db 594 TGAGATCTCTGTCCTGTCATTTGTCATCTTTATAGGAATATCTCTGATGTTGTTTTT 646

RESULT 7 754 bp mRNA linear EST 21-DEC-2000
LOCUS BF669398
DEFINITION 602120693F1 NIH_MGC_56 Homo sapiens cdna clone IMAGE:4277855 5',
BF669398
EST 21-DEC-2000

Query Match	11.4%; Score 494.4; DB 14; Length 581;	
Best Local Similarity	95.9%; Pred. No. 7.3e-52;	
Matches 537; Conservative	0; Mismatches 20; Indels 3; Gaps 3;	
QY 2938	GTGGGTTACTGATGATATAACCAAGAAAAATAATTTAGAGGAATTAAGTTTGA 2997	
Db		
560	GGTGGTACTGATGTTTATACCCAGGAAAAATTTTNGAGGGGATTTAAGTTGGA 501	
QY 2998	AACCACTGTTAA-CAAATCTACCAAGCTGGATTCGCTGCTGATTAAGATGCTGTA 3056	
Db		
500	ACCCACATGTTAACCAANTTCTACCAAGTGGATTTGCCGTGATTAAGATGCTGTA 441	
QY 3057	ACATTTGGG-CACTAGTATATAATTTGAAAAATGTTTATACCA-ATATATAATTTT 3114	
Db		
440	ACATTTGGGCCCACTAGTATAATTTGAAAAATGTTTATACCAATATATAATTTT 381	
QY 3115	TTTAATATACAGTTTCATCAGCTATATTAGTATTTTCAATTAAGTCTAAGATGCCATCAGTG 3174	
Db		
380	TTTAATATACAGTTTCATCAGCTATATTAGTATTTTCAATTAAGTCTAAGATGCCATCAGTG 321	
QY 3175	GTACCAACACCACTGTTTATGCACTGCTAAGAAAGTAAGAGCTGCTGCACTGG 3234	
Db		
320	GTACCAACACCACTGTTTATGCACTGCTAAGAAAGTAAGAGCTGCTGCACTGG 261	
QY 3235	CTCACACCTGTGGAGCCCAAGGAGGAGCATCCTTGAGGCCAGAAAGTTCAAGACCAAC 3294	
Db		
260	CTCACACCTGTGGAGCCCAAGGAGGAGCATCCTTGAGGCCAGAAAGTTCAAGACCAAC 201	
QY 3295	CTGGTCAACATTTGAACACCTGCTCTACAAAAAAGTAAGATTTAGCTGGGT 3354	
Db		
200	CTGGTCAACATTTGAACACCTGCTCTACAAAAAAGTAAGATTTAGCTGGGT 141	
QY 3355	GGGTTGGACATGCTGTAGTCTCAGCTACTCTGAGGCTAAGTGGGAGGATGCTAGA 3414	
Db		
140	GGGTTGGACATGCTGTAGTCTCAGCTACTCTGAGGCTAAGTGGGAGGATGCTAGA 81	
QY 3415	GCCACGGTGTGGAGCTGCAATGAGCTGTGACCAACCACTGGCTCCAGCGTGGGCAA 3474	
Db		
80	GCCACGGTGTGGAGCTGCAATGAGCTGTGACCAACCACTGGCTCCAGCGTGGGCAA 21	
QY 3475	CAGATGAGACCTGTTTCT 3494	
Db		
20	CAGATGAGACCTGTTTCT 1	
RESULT 9		
AA224499		
LOCUS	AA224499 509 bp mRNA linear EST 19-FEB-1997	
DEFINITION	zr16f03.r1 Stratagene NT2 neuronal precursor 937230 Homo sapiens	
	cDNA clone IMAGE:563581 5' similar to TR:G897581 G897581	
	IRON-REGULATORY PROTEIN 2 ; mRNA sequence.	
ACCESSION	AA224499	
VERSION	AA224499.1 GI:1845041	
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
	1 (bases 1 to 509)	
	Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,	
	Chissoe, S., Dietrich, N., DuBuque, T., Favello, A., Gish, W., Hawkins	
	, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore	
	, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T.,	
	Schellenberg, K., Soares, M.B., Tap, F., Thierry-Mieg, J., Trevaskis, E.,	
	Underwood, K., Wohldmann, P., Waterston, R., Wilson, R. and Warrar, M.	
	Generation and analysis of 280,000 human expressed sequence tags	
	Genome Res. 6 (9), 807-828 (1996)	
TITLE		
JOURNAL		
MEDLINE		
COMMENT	Contact: Wilson RK	
	Washington University School of Medicine	
	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108	
	Tel: 314 286 1800	
	Fax: 314 286 1810	

FEATURES	Source	
	1.. 509	/organism="Homo sapiens"
	/db_xref="GDB:5589510"	
	/db_xref="taxon:9606"	
	/clone="IMAGE:663581"	
	/clone_lib="Stratagene NT2 neuronal precursor 937230"	
	/tissue_type="neuroepithelial cells"	
	/dev_stage="Ntera-2 neuroepithelial cells"	
	/lab_host="SOLR (kanamycin resistant)"	
	/note="Organ: brain; Vector: pBluescript SK-; Site: 1:	
	ECORI; Site 2: XhoI; Cloned unidirectionally. primer:	
	Oligo dt. Undirected, exponentially growing neuroepithelial	
	cells (Ntera-2/cl.D1). Average insert size: 1.0 kb;	
	Uni-ZAP XE Vector; -5' adaptor sequence: 5' GAATTCGACGAG	
	3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'."	
BASE COUNT	145 a 104 c 114 g 144 t	2 others
ORIGIN		
Query Match	11.2%; Score 487.2; DB 9; Length 509;	
Best Local Similarity	99.0%; Pred. No. 6e-51;	
Matches 489; Conservative	0; Mismatches 5; Indels 0; Gaps 0;	
QY 36	CCTATTGAAACATTAATGACAGTTTCACATAAGAGTTCTTCGATGATATCTAACTGG 95	
Db		
16	CTTATTGAAACATTAATGACAGTTTCACATAAGAGTTCTTCGATGATATCTAACTGG 75	
QY 96	CACCAAGTATGATGCTTCGCTTACTCAATACGGGCTTTGTGGAGCTCCTGTACGAA 155	
Db		
76	CACCAAGTATGATGCTTCGCTTACTCAATACGGGCTTTGTGGAGCTCCTGTACGAA 135	
QY 156	TTGTGATGGCTTTTAAATGAAGAGGAGATGTTATGAACATTTTAGACTGGAAAAACCA 215	
Db		
136	TTGTGATGGCTTTTAAATGAAGAGGAGATGTTATGAACATTTTAGACTGGAAAAACCA 195	
QY 216	ACAAAGCAATGTGAAGTGCCTTTTCCCTGCGCTGCTTCTTCAAGATTTTAACTGG 275	
Db		
196	ACAAAGCAATGTGAAGTGCCTTTTCCCTGCGCTGCTTCTTCAAGATTTTAACTGG 255	
QY 276	ATACCAAGTATGATGCTTCGCTTACTCAATACGGGCTTTGTGGAGCTCCTGTACGAA 335	
Db		
256	ATACCAAGTATGATGCTTCGCTTACTCAATACGGGCTTTGTGGAGCTCCTGTACGAA 315	
QY 336	TCCTGAGAAAGTCCATCCTCTTGTCCGACAGATCTTACAGTTGACCATTTCTTTACAAAT 395	
Db		
316	TCCTGAGAAAGTCCATCCTCTTGTCCGACAGATCTTACAGTTGACCATTTCTTTACAAAT 375	
QY 396	TCAGTTCAGTAAATGTCATACAGATGCACCAATCCTGAGGCTGACCTGCAGAA 455	
Db		
376	TCAGTTCAGTAAATGTCATACAGATGCACCAATCCTGAGGCTGACCTGCAGAA 435	
QY 456	AGCAGAAAGCTCTCTCCACTTAAAGTGCAGCTTAAGAGCTTCCCTGCAGAGCCAGAC 515	
Db		
436	AGCAGAAAGCTCTCTCCACTTAAAGTGCAGCTTAAGAGCTTCCCTGCAGAGCCAGAC 495	
QY 516	TACCTGCCGAGGAT 529	
Db		
496	TACCTGCCGAGGAT 509	
RESULT 10		
AA104450		
LOCUS	AA104450 567 bp mRNA linear EST 29-OCT-1996	
DEFINITION	mo45a10.r1 Life Tech mouse embryo 15 5dpc 10667012 Mus musculus	
	cDNA clone IMAGE:556506 5' similar to gb:X61147 M.musculus mRNA for	
	iron responsive element binding (MOUSE);, mRNA sequence.	
ACCESSION	AA104450	
VERSION	AA104450.1 GI:1650611	

KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE
AUTHORS Marra, M., Kucuba, T., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Hillier, L., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wyllie, F., Lennon, G., Soares, B., Wilson, R. and Waterston, R.

TITLE The WashU-HHMI Mouse EST Project

JOURNAL Unpublished (1996)

COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LNL ; contact the IMAG Consortium (info@image.llnl.gov) for further information.
MGI:337298
Seq primer: ~28M13 rev1 from Amersham
High quality sequence stop: 466.

FEATURES
source
1..567
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone IMAGE:556506
/clone.lib="Life Tech mouse embryo 15 5dpc 10667012"
/tissue_type="embryo"
/dev_stage="15 5dpc embryos"
/lab_hosts="DH108"
/note="Organ: whole embryo; Vector: pCMV-SPORT2; Site:1; Salt: Site:2; NcBI: Cloned unidirectionally. Primer: Oligo df. 15.5dpc embryos. pCMV-SPORT2 vector."
BASE COUNT 157 a 110 g 150 t
ORIGIN

Query Match 11.2%; Score 485.6; DB 9; Length 567;
Best Local Similarity 92.1%; Pred. No. 8.9e-51;
Matches 523; Conservative 0; Mismatches 44; Indels 1; Gaps 1;

QY 405 TAAATGTGCAATACAGATGCACCAAAATCCTGGAGGTGTGACCTGCAGAAAGCAGGAAA 464
Db 1 TAAATGTGCAATACAGATGCACCAAAATCCTGGAGGTGTGACCTGCAGAAAGCAGGAAA 60

QY 465 GCTCTCTCCACTTAAAGTGCAGCTTACAGAGCTTCCCTGCAGAGGCGCAGACTACCTGCCG 524
Db 61 GCTCTCTCCACTTAAAGTGCAGCTTACAGAGCTTCCCATGTGCGAGGCGCAGACTACCTGCCG 120

QY 525 AGGATCTTGTGATCTCGGAGAACTAGCGCGAACTCAGGAACATTTCTCTCGCAGATTGA 584
Db 121 -GGATCTGTTGATCTCTCGGAGAGCTTAAGCGGAATCTCAGGAACATTTCTCTCAGCAGATTGA 179

QY 585 GAATACACCATCCTGTGTCCTTTTCATTTGCAACCAAGTGCCTTGAACCTTGAACCAAGTGT 644
Db 160 GAATACACCTGCTGTGTGCTTTTCATTTGCAACCAAGTGCCTTGAACCTTGAACCAAGTGT 239

QY 645 AAAAAATCAGAGTAGAATTGCGAGAAATCGAGAGAGCTTCAGTTTTCAGTGAG 704
Db 240 AAAAAATCAGAGTAGAATTGCGAGAAATCGAGAGAGCTTCAGTTTTCAGTGAG 299

QY 705 TTCAGAGTTTAAAGATGTGGCAGTAGTCCCTTCCTGGAACTTGAATGGCTTCATCAAT 764
Db 300 CTCAGGAGCTTTAAAGATGTGGCAGTAGTCCCTTCCTGGAACTTGAATGGCTTCATCAAGT 359

QY 765 AAATCTAGATATTTCTCAAGAGTGGTTTTCGAAAGAAAAGACCTCCTCTTCCAGACAG 824
Db 360 GAATCTAGATATTTATCAAGAGTAGTTTTCGAAAGAAAAGACCTCCTCTTCCAGACAG 419

QY 825 TGTAGTCGCGACAGATTACACATAACGATGGTGAATGGTTTATAGGATTTCTGGGGTGGG 884
Db 420 TGTAGTCGCGACAGATTCTCATATAACCATGTGTGAACGATTTGGGGATTTCTTGGGTGGG 479

QY 885 GGTGTGAGGCAATTAACACAGAGCAGTTATGCTGTGCTTGGCAGTTCCTTCTTACTTTACC 944
Db 480 AGTTGGTGCAATGAGACAGAGGCGAGTTATGCTGTGCTTGGCCTGCCAGTTACTCTTACTTACC 539

QY 945 AGAGGTGTTGGATGTGAGTTAACTGGG 972
Db 540 AGAGGTGTTGGATGTGAGTTAACTGGG 567

RESULT 11
AA313478 498 bp mRNA linear EST 19-APR-1997
LOCUS EST183363 Colon carcinoma (HCC) cell line Homo sapiens cDNA 5' end
DEFINITION similar to iron-responsive element-binding protein, mRNA sequence.
ACCESSION AA313478.1 GI:1965807
VERSION AA313478
KEYWORDS EST
SOURCE human
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 498)
Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A., Bult
, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D., White
, O., Sutton, G., Blake, J.A., Brandon, R.C., Man-Wai, C., Clayton, R.A.,
Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D., Fitzgerald
, L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghagen, N.S., Glodok, A.,
Gnelli, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr., Kelley, J.M.,
Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M.,
Moreno-Palauques, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M.,
Phillips, C.A., Ryder, S.E., Scott, J.L., Saudek, D.M., Shirley, R.,
Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Willy, Y.,
Bednarek, D.P., Cao, L., Cepeda, W.A., Coleman, T.A., Collins, E.J.,
Dimke, P., Feng, D.-F., Ferrie, A., Fischer, C., Hastings, G.A., He, W.W.,
Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Kozak, D.L.,
Kunsch, C., Hungjun, J., Li, H., Meisner, P.S., Olsen, H., Raymond, L.,
Wei, Y.F., Wing, J., Xu, C., Yu, G.L., Ruben, S.M., Dillion, P.J., Fannon
, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M. and
Venter, J.C.
Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
Nature 377 (6547 Suppl), 3-174 (1995)
96026280
Other ESTs: THCB5618
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (<http://www.tigr.org/tldb/hgi/hgi.html>)
Seq primer: M13 Reverse.
FEATURES
Location/Qualifiers
1..498
/organism="Homo sapiens"
/db_xref="ATCC (inhost):109613"
/db_xref="taxon:9606"
/clone.lib="Colon carcinoma (HCC) cell line"
/tissue_type="colon"
/cell_type="KM12C"
/cell_line="KM12C(HCC)-parental human colon carcinoma
;Dukes B2"
/note="Organ: colon; Vector: pBluescript SK-; Site:1;
; EcoRI; Site:2: XhoI"
BASE COUNT 144 a 98 c 113 g 143 t
ORIGIN


```

Query Match      11.1%; Score 482.8; DB 9; Length 498;
Best Local Similarity 99.6%; Pred. No. 2.1e-50;
Matches 484; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGACGCCCAAAAGAGATACGCTTTCAGTACCTTTTAACTTAACTAAATGACAGT 60
DB 13 ATGACGCCCAAAAGAGATACGCTTTCAGTACCTTTTAACTTAACTAAATGACAGT 72

QY 61 TCACATAGAAGTCTTCGATATCTAACTTGGCACCAGATGATGCTTTCGCTTAC 120
DB 73 TCACATAGAAGTCTTCGATATCTAACTTGGCACCAGATGATGCTTTCGCTTAC 132

QY 121 TCAATACGGTCTTGTGGAAGCTCTGTACGAAATTTGATGGCTTTTAAATGAAGA 180
DB 133 TCAATACGGTCTTGTGGAAGCTCTGTACGAAATTTGATGGCTTTTAAATGAAGA 192

QY 181 GAAGTCTTATGAACATTTTACACTGGAACCAACCAACAGCAATCTTGAAGTGCCTTT 240
DB 193 GAAGTCTTATGAACATTTTACACTGGAACCAACCAACAGCAATCTTGAAGTGCCTTT 252

QY 241 TTCCTGCGCGTCTTCTTCAAGATTTTACTGGAATACCAATGATGATTTGCT 300
DB 253 TTCCTGCGCGTCTTCTTCAAGATTTTACTGGAATACCAATGATGATTTGCT 312

QY 301 GCTATGAGGAGGAGCAGTGAAGTCTTGGAGGTGATCCTGAGAAAGTCCATCCTGCTTGT 360
DB 313 GCTATGAGGAGGAGCAGTGAAGTCTTGGAGGTGATCCTGAGAAAGTCCATCCTGCTTGT 372

QY 361 CGACAGATCTTACAGTTCACCATTTTACAAATTCAGTAAATGCAATGCAATACAG 420
DB 373 CGACAGATCTTACAGTTCACCATTTTACAAATTCAGTAAATGCAATGCAATACAG 432

QY 421 AATGACCAATCTCGAGGTGGTACCTGCGAAGCAGAAAGTCTCTCCACTTAA 480
DB 433 AATGACCAATCTCGAGGTGGTACCTGCGAAGCAGAAAGTCTCTCTCACTTAA 492

QY 481 GTGCAG 486
DB 493 GTGCAG 498

RESULT 12
AL121430 467 bp mRNA linear EST 25-FEB-2000
LOCUS Dkfzpf7621247_r1.762 (synonym: lmel2) Homo sapiens cDNA clone
DEFINITION Dkfzpf7621247 5', mRNA sequence.
ACCESSION AL121430
VERSION AL121430.1 GI:5927431
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 467)
AUTHORS Ottenwaelder, B., Obermaier, B., Mewes, H.W., Gassenhuber, J. and
Wiemann, S.
TITLE EST (Ottenwaelder, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: Ottenwaelder B
MIPS Am Klopferplatz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by MedGenomix (Martinsried/Germany) within the cDNA
sequencing consortium of the German Genome Project. No sl sequence
available.
This clone (DKFzpf7621247) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
source location/Qualifiers
1. .467

```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFzpf7621247"
/clone_lib="762 (synonym: lmel2)"
/tissue_type="melanoma (MeWo cell line)"
/dev_stage="adult"
/lab_host="DH10B"
/notice="vector: pSport1; Site_1: NotI; Site_2: SalI"

BASE COUNT 141 a 99 c 110 g 117 t
ORIGIN
Query Match      10.7%; Score 463.8; DB 9; Length 467;
Best Local Similarity 99.6%; Pred. No. 4.8e-48;
Matches 465; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 306 GAGGAGGAGCAGTGAAGTCTTGGAGGTGATCCTGAGAAAGTCCATCTGCTTCTCCGAC 365
DB 1 GAGGAGGAGCAGTGAAGTCTTGGAGGTGATCCTGAGAAAGTCCATCTGCTTCTCCGAC 60

QY 366 AGATCTTACAGTGGACCATCTTTACAAATTTGACTTCAATGTCAGTAAATGTCANATACAGATGC 425
DB 61 AGATCTTACAGTGGACCATCTTTACAAATTTGACTTCAATGTCAGTAAATGTCANATACAGATGC 420

QY 426 ACCAAATCTGGAGGTGGTGCACCTGCAGAAAGCAGGAAAGCTCTCTCCACTTAAAGTGCA 485
DB 121 ACCAAATCTGGAGGTGGTGCACCTGCAGAAAGCAGGAAAGCTCTCTCCACTTAAAGTGCA 480

QY 486 GCCTAAAGAGCTTCCCTGCAGAGGCCAGACTTACCTGCCGAGGATCTTGTGATTTCTGGAGA 545
DB 181 GCCTAAAGAGCTTCCCTGCAGAGGCCAGACTTACCTGCCGAGGATCTTGTGATTTCTGGAGA 240

QY 546 ACTAGGCCCAACTCAGGACATTTCTTCGAGATTTGAGATACACCATCTCTGTGTC 605
DB 241 ACTAGGCCCAACTCAGGACATTTCTTCGAGATTTGAGATACACCATCTCTGTGTC 300

QY 606 TTTTCATTTGCAACAGTGCCTGAAACCTGAAACAGTGTAAAAAATCAAGAGTAGAATT 665
DB 301 TTTTCATTTGCAACAGTGCCTGAAACCTGAAACAGTGTAAAAAATCAAGAGTAGAATT 360

QY 666 CGGAGAAATTCGAGAGAGGCTTCAGTTTTTTAAGTGAGTTCAGAGTTCCTTAAAGAGTGT 725
DB 361 CGGAGAAATTCGAGAGAGGCTTCAGTTTTTTAAGTGAGTTCAGAGTTCCTTAAAGAGTGT 420

QY 726 GGAGTGATCCCTCTCGAAGTGAATGGCTCATCAATAAATAGCTTAG 772
DB 421 GGAGTGATCCCTCTCGAAGTGAATGGCTCATCAATAAATAGCTTAG 467

RESULT 13
BF575050 865 bp mRNA linear EST 12-DEC-2000
LOCUS 602134712F1 NTH_MGC_81 Homo sapiens cDNA clone IMAGE:4289715 5',
DEFINITION mRNA sequence.
ACCESSION BF575050
VERSION BF575050.1 GI:11648762
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 865)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: CLONETECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

```


